

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2006, 18:48:35 ; Search time 544 Seconds

(without alignments)
1615.956 Million cell updates/sec

Title: US-10-697-787-1

Perfect score: 381

Sequence: 1 atggcgtcaactcaagaag.....gttatattttcttggttaa 381

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7673375 seqs, 115364844 residues

Total number of hits satisfying chosen parameters: 15346750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA New:*

- 1: /cgn2_6/ptodata/2/pubpna/US08 NEW PUB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/US06 NEW PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US07 NEW PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/PCr_NEW_PUB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US09 NEW PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/US09 NEW PUB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US10 NEW PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US10 NEW PUB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US11 NEW PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US11 NEW PUB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US11 NEW PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US11 NEW PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	97.8	25.7	517	9	US-11-096-568A-8355	Sequence 8355, Ap
2	70	18.4	876	9	US-11-096-568A-29719	Sequence 29719, A
3	66.8	17.5	957	9	US-11-087-099-102	Sequence 102, App
4	66.2	17.4	896	7	US-10-714-887-273	Sequence 273, App
5	64	16.8	994	9	US-11-096-568A-23894	Sequence 23894, A
6	63.2	16.6	1724	9	US-11-096-568A-4492	Sequence 4492, Ap
7	55.2	14.5	866	7	US-10-714-887-245	Sequence 245, App
8	53.6	14.1	1254	9	US-11-096-568A-21266	Sequence 21266, A
9	42.6	11.2	2891	8	US-10-750-185-38981	Sequence 38981, A
10	42.6	11.2	2891	8	US-10-750-623-38981	Sequence 38981, A
11	40.2	10.6	632	6	US-09-925-065A-664600	Sequence 664600, A
12	38.8	10.2	527	6	US-09-925-065A-525531	Sequence 525531, A
13	38.8	10.2	590	6	US-09-925-065A-496715	Sequence 496715, A
14	38.8	10.2	607	6	US-09-925-065A-496714	Sequence 496714, A
15	38.8	10.2	607	6	US-09-925-065A-496716	Sequence 496716, A
16	38.8	10.2	667	6	US-09-925-065A-944139	Sequence 944139, A
17	38.8	10.2	667	6	US-09-925-065A-944140	Sequence 944140, A
18	38.6	10.1	605	6	US-09-925-065A-580584	Sequence 580584, A
19	38.6	10.1	605	6	US-09-925-065A-580585	Sequence 580585, A
20	38	10.0	759	6	US-09-925-065A-925387	Sequence 925387, A

C 21	37.8	9.9	562	6	US-09-925-065A-416894	Sequence 416894, A
C 22	37.6	9.9	717	9	US-11-096-568A-4827	Sequence 4827, Ap
C 23	37.6	9.9	2208	8	US-10-750-185-26159	Sequence 26159, A
C 24	37.6	9.9	2208	8	US-10-750-623-26159	Sequence 26159, A
C 25	37.4	9.8	687411	7	US-10-330-773-26	Sequence 26, Appl
C 26	37.2	9.8	759	6	US-09-925-065A-925386	Sequence 925386, A
C 27	37	9.7	599	6	US-09-925-065A-387198	Sequence 387198, A
C 28	37	9.7	5455	8	US-10-240-708-33	Sequence 33, Appl
C 29	36.8	9.7	587	6	US-09-925-065A-274001	Sequence 274001, A
C 30	36.8	9.7	588	6	US-09-925-065A-561779	Sequence 561779, A
C 31	36.8	9.7	591	6	US-09-925-065A-312318	Sequence 312318, A
C 32	36.8	9.7	634	6	US-09-925-065A-487340	Sequence 487340, A
C 33	36.8	9.7	634	6	US-09-925-065A-487341	Sequence 487341, A
C 34	36.6	9.6	616	6	US-09-925-065A-694182	Sequence 694182, A
C 35	36.6	9.6	1128	12	US-11-091-883-244	Sequence 244, App
C 36	36.4	9.6	1135	8	US-10-750-185-38542	Sequence 38542, A
C 37	36.4	9.6	1135	8	US-10-750-623-39542	Sequence 39542, A
C 38	36.2	9.5	467	6	US-09-925-065A-226581	Sequence 226581, A
C 39	36.2	9.5	615	6	US-09-925-065A-730683	Sequence 730683, A
C 40	36	9.4	4453	12	US-11-011-332A-99	Sequence 99, Appl
C 41	36	9.4	1125000	8	US-10-995-561-13286	Sequence 13286, A
C 42	35.8	9.4	398	6	US-09-925-065A-217343	Sequence 217343, A
C 43	35.8	9.4	398	6	US-09-925-065A-217344	Sequence 217344, A
C 44	35.8	9.4	472	6	US-09-925-065A-624778	Sequence 624778, A
C 45	35.6	9.3	523	6	US-09-925-065A-201478	Sequence 201478, A

ALIGNMENTS

RESULT 1
US-11-096-568A-8355
; Sequence 8355, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide
; FILE OF INVENTION: Thery
; TITLE OF INVENTION: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 8355
; LENGTH: 517
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(517)
; OTHER INFORMATION: Ceres Seq. ID no. 15225172
US-11-096-568A-8355

Query Match	25.7%	Score 97.8;	DB 9;	Length 517;
Best Local Similarity	69.8%	Pred. No. 6.1e-13;		
Matches 132;	Conservative 0;	Mismatches 57;	Indels 0;	Gaps 0;
Qy	34	TGGACGTTTAGTCAAAACAAGATGTTTCGAGAGGGGCTTGGCAGTTTACGACAGGACACA	93	
Db	90	TGGATCTCGAAGCAGAACAGAGATTTGAGATGCGCTTTCGACAGGACACC	149	
Qy	94	CCCGACCGATGCGACAAATGTGGCAAAGCTGTGGAGGGGAACTGTAGAACAGTGAAG	153	
Db	150	CCAGACAGTGGCACACCGTGGCCAGGCGCTGCGAGGAGGAAACCGTGGAGAGTGAAG	209	
Qy	154	CGCCACTATGACATTTCTCGTCGAGGATCTCATCAACATCGAGACTGGTGTGTCCTTTG	213	
Db	210	AGGCATTATGAGAAGCTCGTGAAGATGTAAGGAGATTGAGGAAGGTACACGTGCCCTC	269	
Qy	214	CCCAATTAC	222	
Db	270	CCCAATTAC	278	

RESULT 2
US-11-096-568A-29719
; Sequence 29719, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 29719
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: misc feature
; LOCATION: (1)..(876)
; OTHER INFORMATION: Ceres Seq. ID no. 4926979
US-11-096-568A-29719

Query Match 18.4%; Score 70; DB 9; Length 876;
Best Local Similarity 60.5%; Pred. No. 1.3e-06;
Matches 115; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 33 ATGGACGTTTGTAGTCAAAACAAGATGTTTCGAGAGGGCCCTTGGCAGTGTTCGACAAAGGACAC 92
|||||
Db 63 ATGGACAGCTGAAGAAAACAAGAAATTCGAAAACGCTTTAGCCTTTTACGACAAAGATAC 122
|||||
QY 93 ACCGACCGATGGCACAATGTGCAAAAGCTGTCGAGGGGAAAACCTGTAGAAAGTGAA 152
|||||
Db 123 TCCCGACAGATGGTCCAGAGTCGCTGCCATGCTTCCCGGTAAAACAGTCGGAGATGTGAT 182
|||||
QY 153 GCGCCACTATGACATCTCGTCGAGGATCTCATCAACATCGAGACTGTCGTGTCCTTT 212
|||||
Db 183 CAACACATACAGAGAGCTTGAGGAGACGTTAGGCAATCCAGAGCTGGTCTTATACCAAT 242
|||||
QY 213 GCCCAATTAC 222
|||||
Db 243 CCCTGGTTAC 252

RESULT 3
US-11-087-099-102
; Sequence 102, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 102
; LENGTH: 957
; TYPE: DNA
; ORGANISM: Glycine max
US-11-087-099-102

Query Match 17.5%; Score 66.8; DB 9; Length 957;
Best Local Similarity 58.6%; Pred. No. 6.9e-06;
Matches 116; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 25 ATCTCACCATTGAGCTTTAGTCAAAAACAAGATGTTTCGAGAGGGCCTTGGCAGTTTACGAC 84
|||||
Db 100 AGCAGAGTGGATAGAGAAGATACAAAGATTTGAAAGTGCCTTGTATATATGAC 159
|||||
QY 85 AAGGACACACCCGACCGATGGCACAATGTGGCAAAAGCTGTCGAGGGGAAAACCTGTAGAA 144
|||||
Db 160 AAGGACACCCAGATAGATGTTCAAGGTGGCCGCTATGATCCCTGGGAAGACTGTGTTT 219
|||||
QY 145 GAAGTAGAGGCCCACTATGACATCTCTCGTAGGATCTCATCAACATCGAGACTGGTCGT 204
|||||

Db 220 GATGTGATCAAGCAATATAGGAACTGGGAAGAAGATGTGAGTGAATCGAAGCGGCAT 279
|||||
QY 205 GTCCCTTTGCCCAATTAC 222
|||||
Db 280 GTTCCGATTCCCGGCTAC 297
|||||

RESULT 4
US-10-714-887-273
; Sequence 273, Application US/10714887
; Publication No. US20060015972A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: HEARD, Jacqueline
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: CREELMAN, Robert
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: CANALES, Roger
; APPLICANT: REPETTI, Peter
; APPLICANT: KUMIMOTO, Roderick W
; APPLICANT: GUTTERSON, Neal
; APPLICANT: REUBER, T. Lynne
; APPLICANT: PINEDA, Omaira
; APPLICANT: SHERMAN, Bradley K
; TITLE OF INVENTION: PLANT TRANSCRIPTIONAL REGULATORS OF DROUGHT STRESS
; FILE REFERENCE: MBI0058-CIP
; CURRENT APPLICATION NUMBER: US/10/714,887
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: 10/412,699
; PRIOR FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/135,134
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 430
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 273
; LENGTH: 896
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G1634; predicted polypeptide sequence is paralogous to G2701
US-10-714-887-273

Query Match 17.4%; Score 66.2; DB 7; Length 896;
Best Local Similarity 58.3%; Pred. No. 9.3e-06;
Matches 116; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 24 AATCTCACCATTGAGCCTTTAGTCAAAAACAAGATGTTTCGAGAGGGCCTTGGCAGTTTACGA 83
|||||
Db 102 AGCTCGAGCTGGACTAAAGACAGACAAAGATTTGAGCGAGCTCTTGTCTCTACGC 161
|||||
QY 84 CAAGGACACACCCGACCGATGGCACAATGTGGCAAAAGCTGTCGAGGGGAAAACCTGTAGA 143
|||||
Db 162 TGATGACACGCCTGATCGCTGGTTCAAAGTTGCTGTATGATCCCTGGAAAGACCATATC 221
|||||
QY 144 AGAAGTAGAGGCCCACTATGACATCTCTCGTAGGATCTCATCAACATCGAGACTGGTCG 203
|||||

Db 222 AGATGTCATGAGCAATCTTAAGCTTGAAGAAGACCTCTTCGATATCGAAGCAGACT 281
QY 204 TGTCCCTTTGCCCAATTAC 222
Db 282 TGTCCCGATCCCGGGTTAC 300

RESULT 5

US-11-096-568A-23894
; Sequence 23894, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 23894
; LENGTH: 994
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(994)
; OTHER INFORMATION: Ceres Seq. ID no. 12416292
US-11-096-568A-23894

Query Match 16.8%; Score 64; DB 9; Length 994;
Best Local Similarity 60.2%; Pred. No. 3e-05;
Matches 106; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 32 CATGGAGTTTGTCAAAACAAGATGTTTCGAGAGGCCTTGGCAGTTTACGACAAGGACA 91
Db 202 CGTGGAGCGCGGAGGAGAACAGCTGTTTCGAGAAGGCACTGGCGCAGATCGACCGGAACG 261
QY 92 CACCCGACGATGGCAATGTGGCAAAAGCTGTGCGAGGGAAACTGTAGAAGAGTGA 151
Db 262 CGCCGACAGTGGGAGAGTGGCGCGTGTGCTTGGAGACGGTTCGACGACGTGA 321
QY 152 AGCGCATATGACATTTCTGTCGAGGATCTCATCAACATCGAGACTGGTCTGTC 207
Db 322 GGAGCCAGTACACCGCTTCGAGAGGACGTGGCGCTCATCGAGGCGCGGGCTC 377

RESULT 6

US-11-096-568A-4492
; Sequence 4492, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 4492
; LENGTH: 1724
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1724)
; OTHER INFORMATION: Ceres Seq. ID no. 13638441
US-11-096-568A-4492

Query Match 16.6%; Score 63.2; DB 9; Length 1724;
Best Local Similarity 56.9%; Pred. No. 5.1e-05;
Matches 116; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 25 ATCTCACCATGGACGTTTGTAGTCAAAACAAGATGTTTCGAGAGGCCTTGGCAGTTTACGAC 84
Db 572 AGCACAATATGACCTCTGAAGAGAACAAGCTCTTTGAAATGCTCTTGCAGTGCATGAT 631
QY 85 AAGGACACACCCGACCGCATGGCACAATGTGGCAAAAGCTGTGCGAGGGGAAAACCTGTAGAA 144
Db 632 AAGGACACACCCGATCGGTGGCACAGATGGCTGAGATATTCCTGGAAAGACAGTGGTT 691
QY 145 GAAGTGAAGCGGCACCTATGACATTCCTGTCGAGGATCTCATCAACATCGAGACTGGTCGT 204
Db 692 GATGTGATAAGGCAGTACCAAGGATTTGGAAGTAGATGTTAGCAATATAGAAGCTGGTTG 751
QY 205 GTCCCTTTGCCCAATTACAGACC 228
Db 752 ATTCCAGTTCCTGGCTATAGTAGC 775

RESULT 7

US-10-714-887-245
; Sequence 245, Application US/10714887
; Publication No. US20060015972A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: HEARD, Jacqueline
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: CREELMAN, Robert
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: CANALES, Roger
; APPLICANT: REPETTI, Peter
; APPLICANT: KUMIMOTO, Roderick W
; APPLICANT: GUTTERSON, Neal
; APPLICANT: REUBER, T. Lynne
; APPLICANT: PINEDA, Omaira
; APPLICANT: SHERMAN, Bradley K
; TITLE OF INVENTION: PLANT TRANSCRIPTIONAL REGULATORS OF DROUGHT STRESS
; FILE REFERENCE: MB10058-CIP
; CURRENT APPLICATION NUMBER: US/10/714,887
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: 10/412,699
; PRIOR FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/135,134
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 430
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 245
; LENGTH: 866
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G2701 reference sequence; predicted polypeptide sequence is par

US-10-714-887-245
Query Match 14.5%; Score 55.2; DB 7; Length 866;
Best Local Similarity 55.9%; Pred. No. 0.0028;
Matches 105; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

```
QY 34 TGGACGTTTAGTCAAAACAAGATGTTTCGAGAGGGCCTTGGCAGCTTTACGACAAGGACACA 93
|||
Db 142 TGGACTAAAGAGAGAAAGATGTTTCGAACGAGCTCTTGGCGATATACGCTGAAGACTCG 201
|||
QY 94 CCCGACCGATGGCACAATGCGCAAAAGCTGTCGGAGGGAAACTGTAGAAGAGTGAAG 153
|||
Db 202 CCTGATCGCTGGTTTAAAGTTGCTTCATGATCCCTGGAAAGACTGTTTTTGTGATGTATG 261
|||
QY 154 CGCCACTATGACATTTCTCGTCGAGGATCTCATCAACATCGAGACTCGTGTGTCCTTTG 213
|||
Db 262 AAGCAATATAGTAGCTTGAAGAGACGTTTTCGATATTGAACGAGGACGTGTTCCTCAIT 321
|||
QY 214 CCCAATTA 221
|||
Db 322 CCTGGTTA 329
|||

RESULT 8
US-11-096-568A-21266
; Sequence 21266, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 21266
; LENGTH: 1254
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1254)
; OTHER INFORMATION: Ceres Seq. ID no. 12402516
US-11-096-568A-21266

Query Match 14.1%; Score 53.6; DB 9; Length 1254;
Best Local Similarity 59.0%; Pred. No. 0.007;
Matches 92; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 82 GACAAGGACACACCCGACCGATGSCACAATGTGGCAAAAGCTGTGCGAGGGGAAAACGTGTA 141
|||
Db 259 GAGGAGGACGGGACGCGAGGTGGGAGAAGCTAGCGGAGGCGCGTCGAGGGGGAAGACGCCG 318
|||
QY 142 GAAGAAGTGAAGCCGCACTATGACATCTCGTCGAGGATCTCATCAACATCGAGACTGGT 201
|||
Db 319 GAGGAGGTGAGGCGGCACACTACGAGCTGCTGGTGAGGACGTCGACGGCATCGAGTCGGGC 378
|||
QY 202 CGTCTCCCTTTGCCCAATTACAAGACCTTTCGAATCT 237
|||
Db 379 CGCTCCGCTTCCGGGTACGCGGCTGACGGCGCT 414
|||

RESULT 9
US-10-750-185-38981
; Sequence 38981, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482

; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 38981
; LENGTH: 2891
; TYPE: DNA
; ORGANISM: Bovine 19866881325079
US-10-750-623-38981

Query Match 11.2%; Score 42.6; DB 8; Length 2891;
Best Local Similarity 56.9%; Pred. No. 2.6;
Matches 78; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 239 ACTCAAGAAGCATCAATGACTTTGACACAAGGTATATACTAAATATCTATATATGATGC 298
|||
Db 1995 ACTGAATATGTTTCATTTGATGATGTAACAGGCAATTAACATAAGTCATATTTTCAGCCAG 2054
|||
QY 299 TCTCGATATATTTTGATAATCATTTAGTCAATTTTGAGAAATCTCTCAAAAAGTTCTTG 358
|||
Db 2055 ACTCTAAATATTTCTAATAATCAATTTTATAAAATTTATCTTCTGAGTTGAAATTTTTTG 2114
|||
QY 359 TAAGTTATATTTCTTTG 375
|||
Db 2115 TGTACTATATTAATGCTG 2131
|||

RESULT 10
US-10-750-623-38981
; Sequence 38981, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 38981
; LENGTH: 2891
; TYPE: DNA
; ORGANISM: Bovine 19866881325079
US-10-750-623-38981

Query Match 11.2%; Score 42.6; DB 8; Length 2891;
Best Local Similarity 56.9%; Pred. No. 2.6;
Matches 78; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 239 ACTCAAGAAGCATCAATGACTTTGACACAAGGTATATACTAAATATCTATATATGATGC 298
|||
Db 1995 ACTGAATATGTTTCATTTGATGATGTAACAGGCAATTAACATAAGTCATATTTTCAGCCAG 2054
|||
QY 299 TCTCGATATATTTTGATAATCATTTAGTCAATTTTGAGAAATCTCTCAAAAAGTTCTTG 358
|||
Db 2055 ACTCTAAATATTTCTAATAATCAATTTTATAAAATTTATCTTCTGAGTTGAAATTTTTTG 2114
|||
QY 359 TAAGTTATATTTCTTTG 375
|||
Db 2115 TGTACTATATTAATGCTG 2131
|||

RESULT 11
US-09-925-065A-664600/c
; Sequence 664600, Application US/09925065A
; Publication No. US20040181048A1
```



```

; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 664600
; LENGTH: 632
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-664600

Query Match      10.6%; Score 40.2; DB 6; Length 632;
Best Local Similarity 58.0%; Pred. No. 6.3;
Matches 69; Conservative 1; Mismatches 49; Indels 0; Gaps 0;

QY 242 CAAGAAGCATCAATGACTTTGACACAAAGGTATATACTAAATCTATATATATGCTCT 301
DB 124 CAAAATATACATGTTTTCACATACAAATATGTTTGAATATGATGATGTTGGAAT 65
QY 302 CGATATATTTGATATATCTTAGTGATTTTGAGAAATCTCTCAAAAAGTCTTGTA 360
DB 64 GCGTAAATGAGATAATTAACATGGTTTGTGATAAGAACATACAAAATCCACTTTTA 6

RESULT 12
US-09-925-065A-525531/c
; Sequence 525531, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 525531
; LENGTH: 527
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-525531

Query Match      10.2%; Score 38.8; DB 6; Length 527;
Best Local Similarity 56.5%; Pred. No. 13;
Matches 70; Conservative 1; Mismatches 53; Indels 0; Gaps 0;

QY 251 TCAATGACTTTGACACAAAGGTATATACTAAATATCTATATATGCTCTCGATATAT 310

```

```

DB 258 TAAAGACTATATATAATAATATATATAATAATATATAATAATATATAATA 199
QY 311 TTGATAATCATCTCTAGTGAATTTTGACAAATCTCTCAAAAAGTCTCTGTAAGTATATTT 370
DB 198 TATATAAGCTGGGCAATCTTAAGGAAATTTCAACACCAAAAAGTTGTGCCAGTTAAATTA 139
QY 371 CTTT 374
DB 138 TCTT 135

RESULT 13
US-09-925-065A-496715/c
; Sequence 496715, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 496715
; LENGTH: 590
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-496715

Query Match      10.2%; Score 38.8; DB 6; Length 590;
Best Local Similarity 58.8%; Pred. No. 13;
Matches 67; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 248 GCATCAATGACTTTGACACAAAGGTATATACTAAATATCTATATATGCTCTCGATAT 307
DB 146 GCATATATTATCATGATACAAATTTATGTATATACAAATATTTTATATCTTCATTCAGTAT 87
QY 308 ATTTTGATAATCATCTAGTGATTTTGAGAAATCTCTCAAAAAGTCTCTGTAA 361
DB 86 AATTGGTTTGATCCTTGATTTTATTTATGCAATTTAAANAATATATTGTGA 33

RESULT 14
US-09-925-065A-496714/c
; Sequence 496714, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846

```

```
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 496714
; LENGTH: 607
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-496714

Query Match      10.2%; Score 38.8; DB 6; Length 607;
Best Local Similarity 58.8%; Pred. No. 13;
Matches 67; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 248 GCATCAATGCTTTGACACAGGTATATACTAAATATCTATATATGATGCTCTCGATAT 307
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 163 GCATATATTATCATGATACAAATTTATGTTATACAATATTTTATATCTTCATTTCAGTAT 104
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 308 ATTTTGATATCATCTAGTGAATTTTGAGAAATCTCTCAAAAAGTTCTTGTA 361
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103 AATTGGTTTGTCACTCTGTATTTTATTTATGCAATTTAAATAATATTATTGTGA 50
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 15
US-09-925-065A-496716/c
; Sequence 496716, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925.065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 496716
; LENGTH: 607
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-496716
```

```
Query Match      10.2%; Score 38.8; DB 6; Length 607;
Best Local Similarity 58.8%; Pred. No. 13;
Matches 67; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 248 GCATCAATGCTTTGACACAGGTATATACTAAATATCTATATATGATGCTCTCGATAT 307
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 163 GCATATATTATCATGATACAAATTTATGTTATACAATATTTTATATCTTCATTTCAGTAT 104
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 308 ATTTTGATATCATCTAGTGAATTTTGAGAAATCTCTCAAAAAGTTCTTGTA 361
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103 AATTGGTTTGTCACTCTGTATTTTATTTATGCAATTTAAATAATATTATTGTGA 50
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: March 10, 2006, 19:35:28
Job time : 546 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2006, 18:46:11 ; Search time 573 Seconds
(without alignments)
5498.487 Million cell updates/sec

Title: US-10-697-787-1

Perfect score: 381

Sequence: 1 atggcgctcaactcaagaag.....gttatattcttggtttaa 381

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA Main:

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	381	100.0	381	8	US-10-697-787-1	Sequence 1, Appli
2	381	100.0	381	9	US-10-512-600-1	Sequence 1, Appli
3	120.6	31.7	760	8	US-10-767-795-270	Sequence 270, App
c	119	31.2	543	7	US-10-021-323-16167	Sequence 16167, A
4	119	31.2	779	8	US-10-767-795-271	Sequence 271, App
5	116.8	30.7	678	7	US-10-424-599-6644	Sequence 6644, Ap
6	110.8	29.1	704	7	US-10-425-114-14832	Sequence 14832, A
7	110.8	29.1	730	7	US-10-424-598-135729	Sequence 135729, A
c	110	28.9	466	7	US-10-260-238-4786	Sequence 4786, Ap
8	108.4	28.5	577	7	US-10-021-323-13836	Sequence 13836, A
c	107.4	28.2	505	7	US-10-021-323-16005	Sequence 16005, A
9	107.4	28.2	532	7	US-10-021-323-14482	Sequence 14482, A
10	107.4	28.2	538	7	US-10-021-323-11476	Sequence 11476, A
c	107.4	28.2	542	7	US-10-021-323-11563	Sequence 11563, A
11	107.4	28.2	604	8	US-10-767-795-5000	Sequence 5000, Ap
12	106	27.8	679	7	US-10-424-599-103390	Sequence 103390, A
c	104.8	27.5	556	3	US-09-770-152-473	Sequence 473, App
13	104.8	27.5	578	6	US-10-225-066A-359	Sequence 359, App
14	104.8	27.5	578	7	US-10-374-780A-2719	Sequence 2719, Ap
15	104.8	27.5	578	9	US-10-225-066A-359	Sequence 359, App
16	103.8	27.2	636	8	US-10-767-795-5934	Sequence 5934, Ap
17	103.4	27.1	584	7	US-10-021-323-9298	Sequence 9298, Ap
18	98.6	25.9	507	7	US-10-424-599-5904	Sequence 5904, Ap

ALIGNMENTS

RESULT 1

US-10-697-787-1
; Sequence 1, Application US/10697787
; Publication No. US20040216182A1
; GENERAL INFORMATION:
; APPLICANT: Agronomics, LLC
; TITLE OF INVENTION: Generation of Plants with Improved Pathogen Resistance and
; FILE REFERENCE: Drought Tolerance
; CURRENT APPLICATION NUMBER: US/10/697,787
; PRIOR FILING DATE: 2003-10-29
; PRIOR APPLICATION NUMBER: US 60/375,333
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: PCT/US03/12981
; PRIOR FILING DATE: 2003-04-24
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-697-787-1

Query Match	100.0%	Score 381;	DB 8;	Length 381;
Best Local Similarity	100.0%	Pred. No. 2.8e-94;		
Matches 381;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGGCGTCAAACTCAAGAAGTTCAATCTCACCATCGAGCTTTAGTCAAAACAAGATGTTTC	60	
Db	1	ATGGCGTCAAACTCAAGAAGTTCAATCTCACCATCGAGCTTTAGTCAAAACAAGATGTTTC	60	
Qy	61	GAGAGGGCTTGGCAGTTTACGACAAAGGACACCCGACCGATGGCACAATGTGGCAAA	120	
Db	61	GAGAGGGCTTGGCAGTTTACGACAAAGGACACCCGACCGATGGCACAATGTGGCAAA	120	
Qy	121	GCTGTGGAGGAAACTGTAGAGAGTGAAGCGCCACTATGACATTCCTCGTCAGGAT	180	
Db	121	GCTGTGGAGGAAACTGTAGAGAGTGAAGCGCCACTATGACATTCCTCGTCAGGAT	180	
Qy	181	CTCATCAACATCGAGACTGGTGGTCCCTTTGCCCAATTACAAGACCTTCGAATCTTAC	240	
Db	181	CTCATCAACATCGAGACTGGTGGTCCCTTTGCCCAATTACAAGACCTTCGAATCTTAC	240	
Qy	241	TCAAGAAGCATCAATGACTTTTGACACAAGGATATATACTATATATATGATGCTC	300	
Db	241	TCAAGAAGCATCAATGACTTTTGACACAAGGATATATACTATATATATGATGCTC	300	

QY 301 TCGATATATTTTCATATCAATCTAGTAGATTTTGAGAAATCTCTCAAAAAGTTCTTGTA 360
Db 301 TCGATATATTTTCATATCAATCTAGTAGATTTTGAGAAATCTCTCAAAAAGTTCTTGTA 360
QY 361 AGTTATATTTCTTTGGTTTAA 381
Db 361 AGTTATATTTCTTTGGTTTAA 381

RESULT 2
US-10-512-600-1
; Sequence 1, Application US/10512600
; Publication No. US20050210546A1
; GENERAL INFORMATION:
; APPLICANT: AGRINOMICS, LLC
; TITLE OF INVENTION: GENERATION OF PLANTS WITH IMPROVED PATHOGEN RESISTANCE
; FILE REFERENCE: AG03-033C-US
; CURRENT APPLICATION NUMBER: US/10/512,600
; CURRENT FILING DATE: 2004-10-22
; PRIOR APPLICATION NUMBER: US 60/375,333
; PRIOR FILING DATE: 2002-04-24
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 1
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-512-600-1

Query Match 100.0%; Score 381; DB 9; Length 381;
Best Local Similarity 100.0%; Pred. No. 2.8e-94; Indels 0; Gaps 0;
Matches 381; Conservative 0; Mismatches 0;

QY 1 ATGGCGTCAAACTCAAGAAGTTCAATCTCACCATGGACGTTTGTAGTCAAAAACAAGATGTTTC 60
Db 1 ATGGCGTCAAACTCAAGAAGTTCAATCTCACCATGGACGTTTGTAGTCAAAAACAAGATGTTTC 60

QY 61 GAGAGGCGCTTGGCAGTTTACGACAAAGGACACACCCGACCGATGGCACAATGTGGCAAAA 120
Db 61 GAGAGGCGCTTGGCAGTTTACGACAAAGGACACACCCGACCGATGGCACAATGTGGCAAAA 120

QY 121 GCTGTCCGAGGAGAACTGTAGAGAGTGAAGCGCCACTATGACATTCCTCGTGGAGAT 180
Db 121 GCTGTCCGAGGAGAACTGTAGAGAGTGAAGCGCCACTATGACATTCCTCGTGGAGAT 180

QY 181 CTATCAACATCGAGACTGCTGTCCTTGGCCAAATACAAAGACCTTCGAATCTAAC 240
Db 181 CTATCAACATCGAGACTGCTGTCCTTGGCCAAATACAAAGACCTTCGAATCTAAC 240

QY 241 TCAAGAGCATCAATGACTTTGACACAAAGGTATATACTAAATATCTATATATGATGCTC 300
Db 241 TCAAGAGCATCAATGACTTTGACACAAAGGTATATACTAAATATCTATATATGATGCTC 300

QY 301 TCGATATATTTTGATATCAATCTAGTAGATTTTGAGAAATCTCTCAAAAAGTTCTTGTA 360
Db 301 TCGATATATTTTGATATCAATCTAGTAGATTTTGAGAAATCTCTCAAAAAGTTCTTGTA 360

QY 361 AGTTATATTTCTTTGGTTTAA 381
Db 361 AGTTATATTTCTTTGGTTTAA 381

RESULT 3
US-10-767-795-270
; Sequence 270, Application US/10767795
; Publication No. US20040181830A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement

FILE REFERENCE: 38-21(53534)B
; CURRENT APPLICATION NUMBER: US/10/767,795
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 117596
; SEQ ID NO 270
; LENGTH: 760
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C42883_2
US-10-767-795-270

Query Match 31.7%; Score 120.6; DB 8; Length 760;
Best Local Similarity 75.4%; Pred. No. 1.5e-22;
Matches 150; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 32 CATGGACGTTTGTAGTCAAAAACAAGATGTTTCGAGAGGCGCTTGGCAGTTTACGACAAGGACA 91
Db 138 CATGGACGCGCAAGCAAAAACAAGATTTTCGAAAGGCGCTTGTAGCTGTTTACGACAAGGACA 197

QY 92 CACCCGACCGATGGCACAATGTGGCAAAAGCTGTGGAGGGGAAAACCTGTAGAAGAAGTGA 151
Db 198 CACCAAGATCGTTGGTACAAATGTTCTAAAGCTGTGGAGGGGAAAACCTGTAGAAGAAGTGA 257

QY 152 AGCGCCACTATGACATTCCTCGTGGAGGATCTCATCAACATCGAGACTGCTGCTGTCCTT 211
Db 258 AGAAGCACTATGAGCTTCTTCTGGAAGATGTTAGACACATCGAGTCGGGTTCCTT 317

QY 212 TGCCCAATTAACAAGACCTT 230
Db 318 CCCCAGACTATTGGACCGT 336

RESULT 4
US-10-021-323-16167/c
; Sequence 16167, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 16167
; LENGTH: 543
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3829-015-Q1-N6-E2
US-10-021-323-16167

Query Match 31.2%; Score 119; DB 7; Length 543;
Best Local Similarity 74.9%; Pred. No. 3.5e-22;
Matches 149; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 32 CATGGACGTTTGTAGTCAAAAACAAGATGTTTCGAGAGGCGCTTGGCAGTTTACGACAAGGACA 91
Db 505 CATGGACGCGCAAGCAAAAACAAGATTTTCGAAAGGCGCTTGTAGCTGTTTACGACAAGGACA 446

QY 92 CACCCGACCGATGGCACAATGTGGCAAAAGCTGTGGAGGGGAAAACCTGTAGAAGAAGTGA 151
Db 445 CACCAAGATCGTTGGTACAAATGTTCTAAAGCTGTGGAGAGAAAACCTGTAGAAGAAGTGA 386

QY 152 AGCGCCACTATGACATTCCTCGTGGAGGATCTCATCAACATCGAGACTGCTGCTGTCCTT 211
Db 385 AGAAGCACTATGAGCTTCTTCTTGAAGATGTTTAGAGCGCATCGAGTCGGGTTCCTT 326

QY 212 TGCCCAATTACAGACCTT 230
| | | | |
Db 325 TCCCGACTATTGGACCGT 307

RESULT 5

US-10-767-795-271
; Sequence 271, Application US/10767795
; Publication No. US20040181830A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53534)B
; CURRENT APPLICATION NUMBER: US/10/767,795
; NUMBER OF SEQ ID NOS: 2004-01-30
; CURRENT FILING DATE: 2004-01-30
; SEQ ID NO 271
; LENGTH: 779
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C42883_1
US-10-767-795-271

Query Match 31.2%; Score 119; DB 8; Length 779;
Best Local Similarity 74.9%; Pred. No. 4.3e-22;
Matches 149; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 32 CATGACGCTTTAGTCAAAACAGATGTTTCGAGAGGCGCTTGGCAGTTTACGACAAGGACA 91
| | | | |
Db 172 CATGACAGCCAGCAAAACAGATTTTCGAAGGCGCTTACGTTTACGACAAGGACA 231
| | | | |
QY 92 CACCCGACGATGGCACAATGTGGCAAAAGCTGTGCGAGGGGAAAACTGTAGAGAAAGTGA 151
| | | | |
Db 232 CACCATGCTGTGTGTACAAATGTTGCTAAAGCTGTGGAGAGAAAACTGTGAGGAAAGTGA 291
| | | | |
QY 152 AGCGCCATGACATTTCTCGTGGAGATCTCATCAATCGAGACTGGTGTGTCCTT 211
| | | | |
Db 292 AGAGCATATGAGCTTCTTCTTGAAGATGTTAGACGCATCGAGTCGGGTGCGGTTCCTT 351
| | | | |

QY 212 TGCCCAATTACAGACCTT 230
| | | | |
Db 352 TCCCGACTATTGGACCGT 370

RESULT 6

US-10-424-599-6644
; Sequence 6644, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 6644
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_106008C.1
US-10-424-599-6644

Query Match 30.7%; Score 116.8; DB 7; Length 678;

Best Local Similarity 70.1%; Pred. No. 1.6e-21;
Matches 157; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
QY 21 TTCAATCTCACCATCGAGCGTTTAGTCAAAACAGATGTTTCGAGAGGCGCTTGGCAGTTTA 80
| | | | |
Db 238 TTCTGACTCTCTTGGACGCCAAACAGAACAGCTGTTTGAAGAAAGCACTTGCACAAATA 297
| | | | |
QY 81 CGACAAGGACACACCCGACCGATGACAAATGTGCAAAAGCTGTGCGAGGGAACACTGT 140
| | | | |
Db 298 TGACAAGGATACCCCTGAGCGCTGGCAGATATGACCAAGCAGTAGGTGGAAAATCTGC 357
| | | | |
QY 141 AGAAGAAGTGAAGCGCCACTATGACATTTCTCGTGGAGGATCTCATCAACATCGAGACTGG 200
| | | | |
Db 358 AGATCAAGTTAAGAGACACTATGAAATCTCTTTGGAGGATCTCAGACACATAGAGTCTGS 417
| | | | |
QY 201 TCGTGTCCCTTTGCCCAATTACAAAGACCTTCGAATCTTAACCTCAA 244
| | | | |
Db 418 CCGTGTCTCTCTTCCCAAGTACAAAGTCCACAGGAAGCAGCACCA 461
| | | | |

RESULT 7

US-10-425-114-14832
; Sequence 14832, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 14832
; LENGTH: 704
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3028-039-A6_FLI
US-10-425-114-14832

Query Match 29.1%; Score 110.8; DB 7; Length 704;
Best Local Similarity 68.1%; Pred. No. 7.3e-20;
Matches 154; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 1 ATGGCGTCAAACTCAAGAAAGTTCAATCTCACCATGCGCGTTTAGTCAAAACAGATGTT 60
| | | | |
Db 150 ATGGCATCCAGTTCAATCTCAGCCTCTGGCTCATGGAGTGTAAAGGACAAAGGCTTT 209
| | | | |
QY 61 GAGAGGCGCTTGGCAGTTTACGACAAGGACACACCCGACCGATGGCACAATGTGGCAAAA 120
| | | | |
Db 210 GAAAAGGCTTTAGCTGTTTATGACAAGGACACTCTGACCGTGTGTTGTTCTCAT 269
| | | | |
QY 121 GCTGTGCGAGGAAAACTGTAGAGAAGTGAAGCGCCACTATGACATTTCTGTCAGGAT 180
| | | | |
Db 270 GCTGTGTGTCGCAAACTCCAGAGGAAGTGAAGAAAGGCACTACGAACTCTCTTTTTCAGGAT 329
| | | | |
QY 181 CTCATCAACATCGAGACTGGTGTGTCCTTTTGGCCCAATTACAAGA 226
| | | | |
Db 330 GTTAAGCATATTGAGTCTGGACGTTGCCATTCCCAAAATTACAAGA 375
| | | | |

RESULT 8

US-10-424-599-135729
; Sequence 135729, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K

```
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 135729
; LENGTH: 730
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_93572C.1
US-10-424-599-135729
```

```
Query Match          29.1%; Score 110.8; DB 7; Length 730;
Best Local Similarity 68.1%; Pred. No. 7.4e-20; Indels 0; Gaps 0;
Matches 154; Conservative 0; Mismatches 72;
```

```
QY 1 ATGGCGTCAAACTCAAGAGTTCAATCTCACCATGGACGTTTGTAGTCAAAAACAAGATGTTTC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 147 ATGGCATCCAGTTCAATCTCAGCCTCTGGCTCATGGAGTGTAAAGGACAACAGGCGCTTT 206
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GAGAGGCGCTTGGCAGTTTACGACAAGGACACACCCGACCGATGGCACAATGTGGCAAAA 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 207 GAAAAGCGCTTGTAGTCTGTTTATGACAAGGACACTCTCGACCGTGTGGTACAAATGTTGCTCAT 266
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 GCTGTGGGAGGAAACTGTAGAGAGTGAAGGCGCACCTATCATCTCTCGTCGAGGAT 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 267 GCTGTGGTGGCAAACTCCAGAGGAGTGAAGGCGCACCTATCATCTCTCTGTTGTCAGGAT 326
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 CTCATCAACATCGAGACTGTCGTGTCCTCTTGGCCCAATTACAGA 226
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 327 GTTAAGCATATTGAGTCTGACGCTGCCATTGCCAAATTACAGA 372
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 9
US-10-260-238-4786/c
; Sequence 4786, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.
; APPLICANT: Moughamer, Todd G.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiyaki
; APPLICANT: Kreps, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 4786
; LENGTH: 466
; TYPE: DNA
; ORGANISM: Musa acuminata
US-10-260-238-4786
```

```
Query Match          28.9%; Score 110; DB 7; Length 466;
Best Local Similarity 67.4%; Pred. No. 9.7e-20; Indels 0; Gaps 0;
Matches 155; Conservative 0; Mismatches 75;
```

```
QY 25 ATCTCACCATGACGTTTGTAGTCAAAACAAGATGTTTCGAGAGGGCCTTGGCAGTTTACGAC 84
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 407 AGCTCGTGTGGACCGCAGCAGACAAGATGTTTCGAGAAAGCCCTCGCGGTGTACGAC 348
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 85 AAGGACACACCCGACCGATGGCACAATGTGGCAAAAGCTGTGGAGGGGAAAACGTGTAGAA 144
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 347 AAGGACACCCGACCGCTGGCACAACGTGGCGCGCGCTCGGCGGCAAGTCAGCGGAA 288
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 145 GAAGTGAAGCGCCACTATGACATTTCTGTCGAGGATCTCATCAACATCGAGACTGTCGT 204
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 287 GAAGTGAAGCGCCACTACGAGCTGTGTGGCGGACATTATCTCATCGAAGAGGCCAG 228
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 205 GTCCCTTTGCCAAATTACAAAGACCTTCGAATCTTAACCTAAGAAGCATCAA 254
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 227 ATGCCTCGAGCCAATTACCGCTCTCTCGGCCACAGGGGATGAAGTGTCSA 178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 10

```
US-10-021-323-13836
; Sequence 13836, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 13836
; LENGTH: 577
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3829-015-Q1-K6-E2
US-10-021-323-13836
```

```
Query Match          28.5%; Score 108.4; DB 7; Length 577;
Best Local Similarity 68.0%; Pred. No. 3e-19; Indels 0; Gaps 0;
Matches 151; Conservative 0; Mismatches 71;
```

```
QY 9 AACTCAAGAAGTTCAATCTCACCATGGACGTTTGTAGTCAAAAACAAGATGTTTCGAGAGGGC 68
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 89 AATGTCATCGATTTCAATGTCTGGTTTCATGGACAGCAACAACAAGATTGCAAGGGC 148
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 69 CTTGGCAGTTTACGACAAGGACACACCCGACCGATGGCACAATGTGGCAAAAGCTGTCCG 128
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 149 TTTAGCTGTTTACGACAAGGACACACACAGATCGTTGGTACAAATGTTGCTTAAAGCTGTGGG 208
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 129 AGGAAAACCTGTAGAAGAAGTGAAGCGCCACTATGACATTTCTGTCGAGAGTTCATCAA 188
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 209 AGAGAAAACCTGTTGAGGAAGTGAAGAAGCACTATGAGCTTCTTCTTGAAGATGTTAGACG 268
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 189 CATCGAGACTGTCGTGTCCTTTGCCCAATTACAAGACCTT 230
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 269 CATCGAGTCGGTCTGGGTTCTTTTCCCGGACTATTGACCGT 310
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 11

```
US-10-021-323-16005/c
; Sequence 16005, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
```

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 16005
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3829-034-Q1-N6-A10
US-10-021-323-16005

Query Match 28.2%; Score 107.4; DB 7; Length 505;
Best Local Similarity 70.2%; Pred. No. 5.4e-19;
Matches 144; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 22 TCAATCTCACCATGGACGTTTAGTCAAAACAAGATGTTCCGAGAGGCGCTTGGCAGTTTAC 81
DB 466 TCAGGTTCGAATGGAGCCCAAGCAAAACAAGTTGTTGAGAACGCTTTAGCTATCTAC 407
QY 82 GACAAGGACACACCCGACCGATGGCAAAATGTGGCAAAAGCTGTCCGAGGGAAGAACTGTA 141
DB 406 GACAAGGATACACCAAGCCGTTGGCACAAGTAGCCAGGCGTGTGGGGGGAAGACCGTG 347
QY 142 GAAGAAGTGAAGGCCCACTATGACATTCCTCGAGGATCTCATCAACATCGAGACTGGT 201
DB 346 GAGGAAGTGAAGTTGCATTACCAGAACCTTGTGGATGACATCAAGCAGATAGAGTCTGGG 287
QY 202 CGTGTCCCTTTGCCCAATTACAAGA 226
DB 286 CACGTGCCTTTGCCCCCTTACAAGA 262

RESULT 12
US-10-021-323-14482
; Sequence 14482, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 14482
; LENGTH: 532
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3829-034-Q1-K6-A10
US-10-021-323-14482

Query Match 28.2%; Score 107.4; DB 7; Length 532;
Best Local Similarity 70.2%; Pred. No. 5.4e-19;
Matches 144; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 22 TCAATCTCACCATGGACGTTTAGTCAAAACAAGATGTTCCGAGAGGCGCTTGGCAGTTTAC 81
DB 56 TCAGGTTCGAATGGAGCCCAAGCAAAACAAGTTGTTGAGAACGCTTTAGCTATCTAC 115
QY 82 GACAAGGACACACCCGACCGATGGCAAAATGTGGCAAAAGCTGTCCGAGGGAAGAACTGTA 141
DB 116 GACAAGGATACACCAAGCCGTTGGCACAAGTAGCCAGGCGTGTGGGGGGAAGACCGTG 175

QY 142 GAAGAAGTGAAGGCCCACTATGACATTCCTCGAGGATCTCATCAACATCGAGACTGGT 201
DB 176 GAGGAAGTGAAGTTGCATTACCAGAACCTTGTGGATGACATCAAGCAGATAGAGTCTGGG 235
QY 202 CGTGTCCCTTTGCCCAATTACAAGA 226
DB 236 CACGTGCCTTTGCCCCCTTACAAGA 260

RESULT 13
US-10-021-323-11476
; Sequence 11476, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 11476
; LENGTH: 538
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3829-002-Q1-K6-E10
US-10-021-323-11476

Query Match 28.2%; Score 107.4; DB 7; Length 538;
Best Local Similarity 70.2%; Pred. No. 5.4e-19;
Matches 144; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 22 TCAATCTCACCATGGACGTTTAGTCAAAACAAGATGTTCCGAGAGGCGCTTGGCAGTTTAC 81
DB 33 TCAGGTTCGAATGGAGCCCAAGCAAAACAAGTTGTTGAGAACGCTTTAGCTATCTAC 92
QY 82 GACAAGGACACACCCGACCGATGGCAAAATGTGGCAAAAGCTGTCCGAGGGAAGAACTGTA 141
DB 93 GACAAGGATACACCAAGCCGTTGGCACAAGTAGCCAGGCGTGTGGGGGGAAGACCGTG 152
QY 142 GAAGAAGTGAAGGCCCACTATGACATTCCTCGAGGATCTCATCAACATCGAGACTGGT 201
DB 153 GAGGAAGTGAAGTTGCATTACCAGAACCTTGTGGATGACATCAAGCAGATAGAGTCTGGG 212
QY 202 CGTGTCCCTTTGCCCAATTACAAGA 226
DB 213 CACGTGCCTTTGCCCCCTTACAAGA 237

RESULT 14
US-10-021-323-11563/c
; Sequence 11563, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880

Job time : 574 secs

```
; SEQ ID NO 11563
; LENGTH: 542
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3829-002-Q1-N6-E10
US-10-021-323-11563

Query Match      28.2%; Score 107.4; DB 7; Length 542;
Best Local Similarity 70.2%; Pred. No. 5.5e-19;
Matches 144; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy  22 TCAATCTCACCATTGGACGTTTAGTCAAAAACAAGATGTTGAGAGAGGCGCTTGGCAGTTTAC 81
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  506 TCAGGTTTGAAGTGGAGCGCCAAAGCAAAACAAGTTGTTGAGAAACGCTTTAGCTATCTAC 447

Qy  82 GACAAGGACACACCCGACCGATGGCACAATGTGGCAAAAGCTGTCGGAGGGAAACTGTA 141
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  446 GACAAGGATACACAGACCGGTGGCACACAGCTAGCCAGGCTGTTGGGGGAAAGACCGTG 387

Qy  142 GAAGAAAGTGAAGCGCCACTATGACATTTCTCGTGAGGATCTCATCAACATCGAGACTGGT 201
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  386 GAGGAAGTGAAGTTGCATTACCAAGACCTTTGTGGATGACATCAAGCAGATAGAGTCTGGG 327

Qy  202 CGTGTCCTTTGCCCAATTACAAGA 226
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  326 CACGTGCCCTTTGCCCCCTTACAAGA 302

RESULT 15
US-10-767-795-5000
; Sequence 5000, Application US/10767795
; Publication No. US2004018130A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53534)B
; CURRENT APPLICATION NUMBER: US/10/767,795
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 117596
; SEQ ID NO 5000
; LENGTH: 604
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C53491_1
US-10-767-795-5000

Query Match      28.2%; Score 107.4; DB 8; Length 604;
Best Local Similarity 70.2%; Pred. No. 5.8e-19;
Matches 144; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy  22 TCAATCTCACCATTGGACGTTTAGTCAAAAACAAGATGTTGAGAGAGGCGCTTGGCAGTTTAC 81
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  99 TCAGGTTTGAAGTGGAGCGCCAAAGCAAAACAAGTTGTTGAGAAACGCTTTAGCTATCTAC 158

Qy  82 GACAAGGACACACCCGACCGATGGCACAATGTGGCAAAAGCTGTCGGAGGGAAACTGTA 141
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  159 GACAAGGATACACAGACCGGTGGCACACAGCTAGCCAGGCTGTTGGGGGAAAGACCGTG 218

Qy  142 GAAGAAAGTGAAGCGCCACTATGACATTTCTCGTGAGGATCTCATCAACATCGAGACTGGT 201
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  219 GAGGAAGTGAAGTTGCATTACCAAGACCTTTGTGGATGACATCAAGCAGATAGAGTCTGGG 278

Qy  202 CGTGTCCTTTGCCCAATTACAAGA 226
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  279 CACGTGCCCTTTGCCCCCTTACAAGA 303

Search completed: March 10, 2006, 20:03:22
```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2006, 18:29:10 ; Search time 2275 Seconds

(without alignments)
9519.716 Million cell updates/sec

Title: US-10-697-787-1

Perfect score: 381

Sequence: 1 atggcgctcaactcaagaag.....gttatattcttggtttaa 381

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sts.*
11: gb_sy.*
12: gb_un.*
13: gb_vi.*
14: gb_htg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	381	100.0	381	15	AY519525	Arabidops
2	381	100.0	68041	15	AC025814	Arabidops
3	273.4	71.8	294	6	CS138004	Sequence
4	151.4	39.7	303	15	AY519524	Arabidops
5	151.4	39.7	334	15	BT005657	Arabidops
6	151.4	39.7	580	15	AK119034	Arabidops
7	149.4	39.2	120977	15	AC025808	Genomic s
8	108.8	28.6	612	15	LES277944	Lycopersi
9	108.8	28.6	3605	15	AJ583670	Lycopersi
10	108.4	28.5	246	15	BT011255	Arabidops
11	108.4	28.5	506	15	BT011070	Arabidops
12	108.4	28.5	100108	14	AP008111	Lotus cor
13	106.6	28.0	97268	15	AP004546	Lotus cor
14	106.6	28.0	102550	14	AP008039	Lotus cor
15	106.6	28.0	108878	14	AP007808	Lotus cor
16	104.8	27.5	306	15	AY519526	Arabidops
17	104.8	27.5	337	15	BT008554	Arabidops
18	104.8	27.5	570	15	BT008698	Arabidops

19	104.8	27.5	578	6	CS137904	Sequence
20	104.8	27.5	54433	15	AC007119	Arabidops
21	103.2	27.1	3213	15	AY954971	Antirrhin
22	102.8	27.0	108844	15	AY517502	Medicago
23	99.4	26.1	512	15	AK112054	Arabidops
24	99	26.0	198354	15	ATAP22	Arabidops
25	99	26.0	198750	15	ATCHR185	Arabidops
26	96.2	25.2	34070	15	AP007251	Oryza sat
27	96.2	25.2	110000	15	AP008207_254	Continuation (255
28	96.2	25.2	110220	14	AP003688	Arabidops
29	96.2	25.2	162297	15	AP004317	Oryza sat
30	92.4	24.3	513	6	CS137972	Sequence
31	92.4	24.3	92861	15	AC006439	Arabidops
32	92	24.1	294	6	CS137908	Sequence
33	92	24.1	294	15	AY519527	Arabidops
34	92	24.1	89035	15	AT122F8	Arabidops
35	92	24.1	198987	15	ATCHR1V90	Arabidops
36	91.6	24.0	91322	14	AP007301	Arabidops
37	89.4	23.5	166907	14	AC146974	Lotus cor
38	84.6	22.2	99587	15	AC007858	Zea mays
39	84.6	22.2	110000	15	AP008211_286	Continuation (287
40	84.6	22.2	156643	15	AC120988	Oryza sat
41	84.4	22.2	108564	14	AC162787	Medicago
42	79.6	20.9	867	15	AY519533	Arabidops
43	79.6	20.9	898	15	BT005473	Arabidops
44	79.6	20.9	1640	15	AK118891	Arabidops
45	79.6	20.9	77758	15	AB016885	Arabidops

ALIGNMENTS

RESULT 1	AY519525	381 bp	mRNA	linear	PLN 07-FEB-2004
LOCUS	Arabidopsis thaliana MYB transcription factor (Atlg75250)	mRNA,			
DEFINITION	complete cds.				
ACCESSION	AY519525				
VERSION	AY519525.1	GI:41618977			
KEYWORDS	Arabidopsis thaliana (thale cress)				
SOURCE	Arabidopsis thaliana				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi				
REFERENCE	1 (bases 1 to 381)				
AUTHORS	Qu,L. and Gu,H.				
TITLE	The MYB Transcription Factor Family in Arabidopsis: A Genome-Wide Cloning and Expression Pattern Analysis				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 381)				
AUTHORS	Qu,L. and Gu,H.				
TITLE	Direct Submission				
JOURNAL	Submitted (07-JAN-2004) Life Sciences, National Laboratory of Protein Engineering and Plant Genetic Engineering, Peking University, Beijing 100871, China				
FEATURES	Location/Qualifiers				
source	1..381				
gene	/organism="Arabidopsis thaliana"				
	/mol_type="mRNA"				
	/db_xref="taxon:3702"				
	1..381				
	/locus_tag="Atlg75250"				
CDS	1..381				
	/locus_tag="Atlg75250"				
	/codon_start=1				
	/product="MYB transcription factor"				
	/protein_id="AA09995_1"				
	/db_xref="GI:41618978"				
	/translation="MASNSRSSISPTWTFSQNMFERALAVYDKDTPDRMHNVAKAVGG KTVSEVKRHYDILVEDINLETGRVLPNFKYPTFESNRSINDFDTRYITKLYLMLSI YFDNHSSDFEKFQSKVLVSISLV"				
ORIGIN					

Query Match 100.0%; Score 381; DB 15; Length 381;
Best Local Similarity 100.0%; Pred. No. 3.4e-84;
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTCAAACTCAAGAAGTTCAATCTCACCATGACGTTTAGTCAAAACAAGATGTTTC 60
DB 1 ATGGCGTCAAACTCAAGAAGTTCAATCTCACCATGACGTTTAGTCAAAACAAGATGTTTC 60
QY 61 GAGAGGCGCTTGGCAGTTTACGAAGGACACACCCGCGATGGCAGCAATGTGGCAAAA 120
DB 61 GAGAGGCGCTTGGCAGTTTACGAAGGACACACCCGCGATGGCAGCAATGTGGCAAAA 120
QY 121 GCTGTGCGAGGGAAGAACTGTAGAAGAGTGAAGCGCCACTATGACATTTCTGTCGAGGAT 180
DB 121 GCTGTGCGAGGGAAGAACTGTAGAAGAGTGAAGCGCCACTATGACATTTCTGTCGAGGAT 180
QY 181 CTCATCAACATCAGACTGGTGTGTCCTTCCCTTCCCAATTAACAGACCTTCGATCTAAC 240
DB 181 CTCATCAACATCAGACTGGTGTGTCCTTCCCTTCCCAATTAACAGACCTTCGATCTAAC 240
QY 241 TCAAGAGCATCAATGACCTTTGACACAGGATGATATTAACATTAATCTATATATGATGCTC 300
DB 241 TCAAGAGCATCAATGACCTTTGACACAGGATGATATTAACATTAATCTATATATGATGCTC 300
QY 301 TCGATATATTTGATATCAATCTAGTGATTTTGAGAAATTTCTCTCAAAAAGTTCTGTGA 360
DB 301 TCGATATATTTGATATCAATCTAGTGATTTTGAGAAATTTCTCTCAAAAAGTTCTGTGA 360
QY 361 AGTTATATTTCTTTGGTTTAA 381
DB 361 AGTTATATTTCTTTGGTTTAA 381

RESULT 2
AC025814
LOCUS Arabidopsis thaliana chromosome 1 BAC F22H5 genomic sequence, complete sequence.
DEFINITION
ACCESSION AC025814
VERSION AC025814.7 GI:12331602
KEYWORDS HTG.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 68041)
Lin, X., Kaul, S., Town, C.D., Benito, M.-I., Creasy, T.H., Haas, B.J., Wu, D., Maiti, R., Ronning, C.M., Koo, H., Fujii, C.Y., Utterback, T.R., Barnstead, M.E., Bowman, C.L., White, O., Nierman, W.C. and Fraser, C.M. Arabidopsis thaliana chromosome 1 BAC F22H5 genomic sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 68041)
Lin, X. and Kaul, S.
JOURNAL Direct Submission
AUTHORS Submitted (15-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, xlin@tigr.org
JOURNAL
REFERENCE 3 (bases 1 to 68041)
Town, C.D. and Kaul, S.
JOURNAL Direct Submission
AUTHORS Submitted (05-OCT-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
JOURNAL
REFERENCE 4 (bases 1 to 68041)
Town, C.D. and Kaul, S.
JOURNAL Direct Submission
AUTHORS Submitted (12-SEP-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
JOURNAL
REFERENCE 5 (bases 1 to 68041)
Town, C.D. and Kaul, S.
JOURNAL Direct Submission
AUTHORS Submitted (12-OCT-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

6 (bases 1 to 68041)
Town, C.D. PhD.
Direct Submission
Submitted (22-JAN-2001)
On Jan 22, 2001 this sequence version replaced gi:12280794.
Address all correspondence to: at@tigr.org

BAC clone F22H5 is from Arabidopsis thaliana chromosome 1.
The orientation of the sequence is from SP6 to T7 end of the BAC clone.

Genes were identified by a combination of several methods: Gene prediction programs including Genscan+ (Chris Burge, http://CCR-081.mit.edu/GENSCAN.html), GeneMarkHM (Mark Borodovsky, http://genemark.biology.gatech.edu/GeneMark/), GlimmerA (a variant of GlimmerM, see Mihaela Pertea, http://www.tigr.org/softlab/glimmerm.htm/glimmerm.html, and GeneSplicer (Mihaela Pertea and Steven Salzberg, contact mpertea@tigr.org), searches of the complete sequence against a peptide database and the plant EST database at TIGR (http://www.tigr.org/tdb/tgi.shtml). Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as unknown proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as hypothetical proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are identified by RepeatMasker (Arian Smit, http://ftp.genome.washington.edu/RM/RepeatMasker.html).

FEATURES
source

1. .68041
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="F22H5"
/ecotype="Columbia"
/complement(418..5404)
/gene="F22H5.15"
complement(join(<418..594,684..773,873..938,1034..1095,1179..1227,1319..4820,5040..5404))
/gene="F22H5.15"
complement(join(<418..594,684..773,873..938,1034..1095,1179..1227,1319..4820,5040..5404))
/codon_start=1
/product="unknown protein, 3' partial; 5120-134"
/protein_id="AAC12690.1"
/db_xref="GI:10092277"
/translation="MEYKSPATTFESRKISNNRSHLSFSANAVYDGVFSSPVNSKS PLVDYGEIFRGGSPSSIPFLDVPENLVGVKVDVRSSKLDYSSVFGGLGACDPAT PKEVITKSEKTSINEDKRNRRKGGSDVPLCNKGKSPKPMVRMKHSDISYHQTVV RNENGATHLTQVPATMPGPIQVVDNTSLHKIESKTPPAVEKKLPCNREEVYK ASRKQSGKTEVPENIFARDGGRDSDCTVSGEYRDVDPKPPSFQCTLNGEHGAS ERLGSLNPSERYETEDSPSPYFADETDENSVAESAALKAIIEA01RWNI AQOMKKKSGFRSCKLSCDSKLENKNTKEGITEBSRDNNSQLLGEMKPSQ SFSGNEGDHAKARKLWGPVEGLKTSDDHPEELBEQDITILEEQARGRKHWLP GTFKSVNNSKQEPENLAPKPEPTKQEVQIPENFPYFQGLGSKLCVVEAF7P SKVSKQDEKQFTKENSTVTVMQDESDSQEMLAGIPVETYLREVEETPQQTESKS EKNIEKSESTICAFTESSONKEGTQVKSACKFEDGSGVKDVGNDGHTCNVD OGEKEIVSEPOEMLVGPDDSKTYVREVEETPTSLNKTQSDSDSFGAMVFNWVISE PGNVIDVSAVHKVPRRRWKTSDYNNIKAPKGNRPWOLESAENETAMFHEE GVRIHVASBEISTGSGAASGLQENWTLVKMFQNFQATDKGDETYCLVESERG HLDIHQAEKYEVEETVRYNVEAYHTRENEDESAQETYCRKEDGKVEVQKTSI VRELIGEELEMALEEDVEQAESEAGVQGLSELNEINEHADHAEMLYDRSETLS NNSRRFDOTQEOABETMIDSGIDTDSRSSFPMQGSQSYIEBVGTEHQHRSDDPPEKA TSVNTEETHEIDSDSIQSGSVVDDDRSLQDQGSQAEKXHEDELEETKEISDEM K TSLGVRNGDKKELEHQFCQENETYSNVVEAESCRFPNGEIIIGAAATNGMKNKE GEEGESCRSMEEBGATSDIGAATDGNMKNENEGESCRSMEEBGATSDISQNKVA ETVEEHLKKIDETREKERQKQVVMVRAIRAFADAFADAMEEGATADAKAIVA HRREVPRKSEKGSVEYNDKLSAEMQAKLRAERAVERAITEVREERAMEKALSGK SAASQAKSYGGKSGFSSGERRGSSSTENKSGPSNSNQTQPTQRCARSRHQ

gene
mRNA
CDS

```
RTSDRAAEALAEKKLRDLTKOETORNRRLABALDADVKRWSSGKNNRLALISTLOY
ILGABSGWKPIPLTDLVSSVRKAYRKATLYVHPDKLQORGASTQOKYICKVFDLL
K"
complement(9354..10899)
/gene="F22H5.16"
/note="similar to GI:1708420 from [Arabidopsis thaliana]
(J. Biol. Chem. 270 (44), 26224-26231 (1995))"
complement(join(<9354..9638,9750..9953,10043..10328,
10461..10595,10671..>10899))
/gene="F22H5.16"
complement(join(9481..9638,9750..9953,10043..10328,
10461..10595,10671..10856))
/gene="F22H5.16"
/codon_start=1
/product="NADPH oxidoreductase, putative; 10572-9197"
/protein_id="AAG12695.1"
/db_xref="GI:10092282"
/translation="MTEKSKILVIGGTGYMGFEIVESAKAGNFTFALVREASLSDP
VSKTIQSKDGLVTLHGDLNDSLVKAIQVDVISTIGHKQIPDQTKIIISAIKE
AGNVRFLPAEFGIDVERTSAVEPAKSFAGKQVIRRAIEAGIPTYVVSNGSAGFY
LRTLQFESGLISHRDRAIFGDKNVPRDKVTILGDNKAVINKEDVAAYMIKA
VDLRLAKTLVISPNNILSNMNVTLWEKKIGKSLKTHISEQILKSIQVPIDVF
KSINHAVVKGQDTSFTIEPWFGEASVLDPVKVTSIDEVLSQFT"
complement(11122..12573)
/gene="F22H5.18"
/note="similar to GI:1708420 from [Arabidopsis thaliana]
(J. Biol. Chem. 270 (44), 26224-26231 (1995))"
complement(join(<11122..11422,11516..11734,11791..12034,
12117..12251,12333..>12573))
/gene="F22H5.18"
complement(join(11235..11422,11516..11734,11791..12034,
12117..12251,12333..12518))
/gene="F22H5.18"
/codon_start=1
/product="NADPH oxidoreductase, putative; 12234-10951"
/protein_id="AAG12680.1"
/db_xref="GI:10092267"
/translation="NASESKILVIGGTGHIGKLIIEASVRAGHSTLALVREASLSDP
NKGTVQNFQDGVTLHGDLNDSLVKAIQADVISTVGSMDLQDQTKIIISAIKE
AGNVRFLPESFGMDVKSASAVEPAKSFAGKLOPRDIEAGIPTYVLTNYPAGY
LPTLVQLRPLTSPPRDKVIFGDKNVKVEYFIIVINKEDIAAYTIKAVDDPRTLNK
TLIINPNNILSNMNVTLWEKKIGKSLKTHISEQILKSIQVPIDVF
VFKGDQNTFIEPFGFEASELYPDIKYTSIDEVLSYFALGTSLNT"
complement(12872..14426)
/gene="F22H5.17"
/note="similar to GI:1708420 from [Arabidopsis thaliana]
(J. Biol. Chem. 270 (44), 26224-26231 (1995))"
complement(join(<12872..13216,13298..13513,13578..13821,
13974..14108,14193..>14426))
/gene="F22H5.17"
complement(join(13053..13216,13298..13513,13578..13821,
13974..14108,14193..14378))
/gene="F22H5.17"
/codon_start=1
/product="NADPH oxidoreductase, putative; 14094-12769"
/protein_id="AAG12677.1"
/db_xref="GI:10092264"
/translation="NATEKSKILVIGGTGYIGKFLVEASAKAGHSTFALVREATLSDP
VKGTVQSKDGLVTLHGDLNDSLVKAIQVDVISTVGSMDLQDQTKIIISAIKE
AGNVRFLPESFGVDVDRTSASAVEPAKSFAGKQIRRTIEAGIPTYVAVTGCFGGY
LPTLVQFEPGTSPPRDKVTILGDNKAEYFAAVINKEDIAAYTIKAVDDPRTLNKI
LVIKPSNNTLSNMTVTLWEKKIGKSLKTHLPEQLKSIQESPIPNVLSINHAV
FVNGDNTISIPESFGFEASELYPDVKVTSIDEVLSYFA"
15081..16153
/gene="F22H5.1"
/note="similar to GI:6939839 from [Oryza sativa]"
join(15081..15228,15380..15470,15661..>16153)
/gene="F22H5.1"
join(15171..15228,15380..15470,15661..16153)
/gene="F22H5.1"
/codon_start=1
/product="GSH-dependent dehydroascorbate reductase 1,
putative; 14887-15869"
```

```
/protein_id="AAG12679.1"
/db_xref="GI:10092266"
/translation="MALDICVAVGAPDVLGDCPPSORVLLTLEKKLPYKTHLINV
SDKPQWFLDISPEGKVPVVKLDGKVVADSDVIVGLLEKYPEPSLKTPEPASVGSKI
FGAFVTFUKSDKNDSEKALVDELEALEHLENLHKTSHSGPFAAGEKITAVDLSLAPKLIH
LEVALGHYKNWSVPESLTSVRNYAKALFSRESFENTKAKKEIVVAGWESKVNA"
complement(16410..18554)
/gene="F22H5.2"
complement(join(<16410..16558,16703..16830,17050..>18554))
/gene="F22H5.2"
complement(join(16410..16558,16703..16830,17050..18554))
/gene="F22H5.2"
/codon_start=1
/product="unknown protein; 18270-16126"
/protein_id="AAG12682.1"
/db_xref="GI:10092269"
/translation="MPPFRFRIPFPSSNSPRISSGTSSPSPPTPPSPRPPRPAGIA
Query Match 100.0%; Score 381; DB 15; Length 68041;
Best Local Similarity 100.0%; Pred. No. 3.4e-84;
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGTCAAACTCAAGAAGTTCAATCTCACCATCGAGCTTTAGTCAAAAACAAGATGTTTC 60
Db 20955 ATGGCGTCAAACTCAAGAAGTTCAATCTCACCATCGAGCTTTAGTCAAAAACAAGATGTTTC 21014
QY 61 GAGAGGGCTTGGCAGTTTACGACAAGGACACACCCGACCGATGCGACAATCTGCGCAAAA 120
Db 21015 GAGAGGGCTTGGCAGTTTACGACAAGGACACACCCGACCGATGCGACAATCTGCGCAAAA 21074
QY 121 GCTGTGCGAGGAGAAAATCTGTAGAAGAGTGAAGCGCCACTATGACATTCCTCGAGGAT 180
Db 21075 GCTGTGCGAGGAGAAAATCTGTAGAAGAGTGAAGCGCCACTATGACATTCCTCGAGGAT 21134
QY 181 CTATCAACATCGAGACTGGTCGTCCTCTTGGCCAAATATCAAGACCTTCGAATCTAAC 240
Db 21135 CTATCAACATCGAGACTGGTCGTCCTCTTGGCCAAATATCAAGACCTTCGAATCTAAC 21194
QY 241 TCAAGAAGCATCAATGACTTTGACACAAGGATATATACTAAATATCTATATGATGCTC 300
Db 21195 TCAAGAAGCATCAATGACTTTGACACAAGGATATATACTAAATATCTATATGATGCTC 21254
QY 301 TCGATATATTTGATAATCAATCTAGTAGTATTGAGAAATTCCTCAAAAAGTTCCTTGTA 360
Db 21255 TCGATATATTTGATAATCAATCTAGTAGTATTGAGAAATTCCTCAAAAAGTTCCTTGTA 21314
QY 361 AGTTATATTTCTTTGGTTTAA 381
Db 21315 AGTTATATTTCTTTGGTTTAA 21335
RESULT 3
CS138004 LOCUS 294 bp DNA linear PAT 09-AUG-2005
DEFINITION Sequence 975 from Patent WO2005047516.
ACCESSION CS138004
VERSION CS138004.1 GI:72058666
KEYWORDS Arabidopsis thaliana (chale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
1
REFERENCE
AUTHORS Heard,J.E.
.. TITLE Plant transcriptional regulators
JOURNAL Patent: WO 2005047516-A 975 26-MAY-2005;
Mendel Biotechnology, Inc. (US)
FEATURES
Location/Qualifiers
1..294
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
```

ORIGIN		/note="G2721"	
Query Match		71.8%; Score 273.4; DB 6; Length 294;	
Best Local Similarity		97.9%; Pred. No. 2.2e-57;	
Matches 277; Conservative		0; Mismatches 6; Indels 0; Gaps 0;	
QY		1 ATGGCGTCAAACTCAAGAAGTTCAATCTCACCATCGACGCTTTAGTCAAAACAAGATGTTTC 60	
DB		1 ATGGCGCTTAGTCTTATGAGCTCGAGCTCTTCTTGACGCTTAAGCAAAACAAGATGTTTC 60	
QY		61 GAGAGGGCCTTGCGAGTTTACGACAAGGACACACCCGACCGATGGCACAAATGTGGCAAAA 120	
DB		61 GAGAGGGCTTTAGCGCTTTACGATAAGACACTCCCGACCGTTGGCAAAACGTTGGCTAAA 120	
QY		121 GCTGTGGAGGAAAACCTGTAGAGAAGTGAAGCGCCACTATGACATTCCTCGTCGAGGAT 180	
DB		121 CGAGTTGGAAGTAAATCTCGAGAGGAAGTTAAACGTCCTACGACATCTCTCGTTGAAGAT 180	
QY		181 CTCATCAACATCGAGACTGGTCTGCTCCCTTTTCCCAATTACAAGACCTTCGA-----A 234	
DB		181 CTCATGAACATCGAACAAGACTTAGTACCTTTTGCCCTAAATACAAAACCGTCGATGTGGA 240	
QY		235 TCTAACTCAAGAGCATCAATGACTTTTGACACAGGTATATATAACTAA 281	
DB		241 AGTAAATCTAGAGCATCGATGATTTTGAGGTTAATGAAGAA 287	
RESULT 5		BT005657 334 bp mRNA linear PLN 15-MAR-2003	
LOCUS		Arabidopsis thaliana clone US1233 putative myb family transcription	
DEFINITION		factor (Atlg19510) mRNA, complete cds.	
ACCESSION		BT005657	
VERSION		BT005657.1 GI:28973504	
KEYWORDS		FLI CDNA.	
SOURCE		Arabidopsis thaliana (thale cress)	
ORGANISM		Arabidopsis thaliana	
REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
AUTHORS		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;	
		rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsia.	
		1 (bases 1 to 334)	
		Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Heuan,V.W., Lee,J.M.,	
		Onodera,C.S., Quach,H.L., Tang,C., Toriumi,M., Wong,C., Wu,H.C.,	
		Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,	
		Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M.,	
		Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P.,	
		Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W.,	
		Ecker,J.R. and Theologis,A.	
TITLE		Arabidopsis Open Reading Frame (ORF) Clones	
JOURNAL		Unpublished	
REFERENCE		2 (bases 1 to 334)	
AUTHORS		Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,	
		Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C.,	
		Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,	
		Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M.,	
		Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P.,	
		Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W.,	
		Ecker,J.R. and Theologis,A.	
TITLE		Direct Submission	
JOURNAL		Submitted (15-MAR-2003) Plant Gene Expression Center, 800 Buchanan	
COMMENT		Street, Albany, CA 94710, USA	
		The RIKEN Genomic Sciences Center (GSC) members carried out the	
		collection and clustering of RAPL cDNAs (RAPL cDNA : RIKEN	
		Arabidopsis Full-Length cDNA): Seki,M., Narusaka,M., Ishida,J.,	
		Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,	
		Hayashizaki,Y. and Shinozaki,K.	
		The Salk, Stanford, PGECC (SSP) Consortium members constructed and	
		sequenced the PUNI (ORF) clones using the RAPL cDNAs: Yamada,K.,	
		Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,	
		Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C.,	
		Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,	
		Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M.,	
		Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P.,	
		Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W.,	
		Ecker,J.R. and Theologis,A.	
ORIGIN		Yamada,K. (SSP/PGECC) and Seki, M. (RIKEN GSC) contributed equally	
		to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP	
		/PGECC) contributed equally to this work as PIs.	

Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.

```

FEATURES
    source
        Location/Qualifiers
            1..334
                /organism="Arabidopsis thaliana"
                /mol_type="mRNA"
                /db_xref="taxon:3702"
                /chromosome="1"
                /clone="U51233"
                /ecotype="Columbia"
                /note="This clone is in pUNI 51."
            1..334
                /gene="Atlg19510"
            1..303
                /gene="Atlg19510"
                /codon_start=1
                /evidence=experimental
                /product="putative myb family transcription factor"
                /protein_id="AA064077.1"
                /db_xref="GI:28973505"
                /translation="MASSSMSSSSWTSKONKMPERALAVYDKTDPDRWQNVAKAVGS
                KSAEVKSHYDLVLEDLMNIEQDLVPLPKYKTVGVSKSRGIDDFDLRLMKMNRQ"
            3'UTR
                304..334
                /gene="Atlg19510"

ORIGIN
    Query Match          39.7%;   Score 151.4;   DB 15;   Length 334;
    Best Local Similarity 73.2%;   Pred. No. 5.5e-27;
    Matches 210;   Conservative 0;   Mismatches 71;   Indels 6;   Gaps 1;

QY  1  ATGGCGTCAAACTCAAGAAGTTCAATCTCACCATGGACGCTTTAGTCAAAACAAGATGTTTC 60
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB  1  ATGGCCTCTAGTTCTATGAGCTCGAGCTTTCTTGGAGCTCTAAGCAAAACAAGATGTTTC 60

QY  61  GAGAGGCGCTTGGCGAGTTTACGACAAAGGACACACCCGACCGATGGCACAAATGTGCAAAA 120
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB  61  GAGAGGCGCTTGGCGAGTTTACGACAAAGGACACACCCGACCGATGGCACAAATGTGCTAAA 120

QY  121  GCTGTCGAGGAGAACTGTAGAAAGTGAAGCCCACTATGACATTCCTCGTCAGGAT 180
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB  121  GCAGTTGGAAGTAAATCTGACAGGAAGTTAAACGTCACGACATCCTCGTTGAAGAT 180

QY  181  CTCATCAACATCGAGACGTGTCGTCCTTTGGCCCAATTACAAGACCTTCGA-----A 234
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB  181  CTCATGAACATCGAACAGACTTAGTACCTTTGGCTTAATACAAAACCGTCGATTTGA 240

QY  235  TCTAACTCAAGAAGCATCAATGACTTTGACACAAGGTATATACTAA 281
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB  241  AGTAAATCTAGAGGCATCGATGATTTGATTTGAGGTTAATGAAGAA 287

RESULT 6
AK119034
LOCUS
DEFINITION
    Arabidopsis thaliana Atlg19510 mRNA for putative myb-related
    protein, complete cds, clone: RAF21-36-F10.
ACCESSION
    AK119034
VERSION
    AK119034.1 GI:26453067
KEYWORDS
    FLI CDNA; CAP trapper.
SOURCE
    Arabidopsis thaliana
    ORGANISM
        Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
        Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
        rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
    1
        Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J.,
        Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P.,
        Kawai, J., Hayashizaki, Y. and Shinozaki, K.
        Arabidopsis thaliana full-length cDNA
        Published Only in Database (2002)
    2 (bases 1 to 580)
        Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J.,
        Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P.,

FEATURES
    source
        Location/Qualifiers
            1..580
                /organism="Arabidopsis thaliana"
                /mol_type="mRNA"
                /db_xref="taxon:3702"
                /chromosome="1"
                /clone="RAF21-36-F10"
                /ecotype="Columbia"
                /note="common name: thale cress"
            1..580
                /gene="Atlg19510/F18014_31"
            88..390
                /gene="Atlg19510/F18014_31"
                /codon_start=1
                /product="putative myb-related protein"
                /protein_id="BAC43610.1"
                /db_xref="GI:26453068"
                /translation="MASSSMSSSSWTSKONKMPERALAVYDKTDPDRWQNVAKAVGS
                KSAEVKSHYDLVLEDLMNIEQDLVPLPKYKTVGVSKSRGIDDFDLRLMKMNRQ"

ORIGIN
    Query Match          39.7%;   Score 151.4;   DB 15;   Length 580;
    Best Local Similarity 73.2%;   Pred. No. 5.5e-27;
    Matches 210;   Conservative 0;   Mismatches 71;   Indels 6;   Gaps 1;

QY  1  ATGGCGTCAAACTCAAGAAGTTCAATCTCACCATGGACGCTTTAGTCAAAACAAGATGTTTC 60
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB  88  ATGGCCTCTAGTTCTATGAGCTCGAGCTTTCTTGGAGCTCTAAGCAAAACAAGATGTTTC 147

QY  61  GAGAGGCGCTTGGCGAGTTTACGACAAAGGACACACCCGACCGATGGCACAAATGTGCAAAA 120
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB  148  GAGAGGCGCTTGGCGAGTTTACGACAAAGGACACACCCGACCGATGGCACAAATGTGCTAAA 207

QY  121  GCTGTCGAGGAGAACTGTAGAAAGTGAAGCCCACTATGACATTCCTCGTCAGGAT 180
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB  208  GCAGTTGGAAGTAAATCTGACAGGAAGTTAAACGTCACGACATCCTCGTTGAAGAT 267

QY  181  CTCATCAACATCGAGACGTGTCGTCCTTTGGCCCAATTACAAGACCTTCGA-----A 234
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB  268  CTCATGAACATCGAACAGACTTAGTACCTTTGGCTTAATACAAAACCGTCGATTTGA 327

QY  235  TCTAACTCAAGAAGCATCAATGACTTTGACACAAGGTATATACTAA 281
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB  328  AGTAAATCTAGAGGCATCGATGATTTGATTTGAGGTTAATGAAGAA 374

RESULT 7
AC025808/c
LOCUS
DEFINITION
    Arabidopsis thaliana thaliana BAC F18014 from
    chromosome I, complete sequence.
ACCESSION
    AC025808
VERSION
    AC025808.8 GI:7636235
KEYWORDS
    HTG.
SOURCE
    Arabidopsis thaliana (thale cress)
    ORGANISM
        Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
        Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

```

Kawai, J., Hayashizaki, Y. and Shinozaki, K.
 Direct Submission
 Submitted (25-NOV-2002) Motoaki Seki, RIKEN Genomic Sciences
 Center; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa
 230-0045, Japan (E-mail: msaki@gs.riken.go.jp,
 URL: http://pfgweb.gsc.riken.go.jp, Tel: 81-45-503-9625,
 Fax: 81-45-503-9586)
 An Arabidopsis full-length cDNA library was constructed essentially
 as reported previously (Seki et al. (1998) Plant J. 15:707-720;
 Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI
 and XhoI was ligated to modified Lambda PLC-1-E vector (Carninci et
 al. (2001) Genomics 77:79-90) digested with BamHI and SalI.
 This clone is in a modified pBluescript vector.
 Please visit our web site (http://pfgweb.gsc.riken.go.jp/) for
 further details.

FEATURES

source

1..580
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="1"
 /clone="RAF21-36-F10"
 /ecotype="Columbia"
 /note="common name: thale cress"

gene

1..580
 /gene="Atlg19510/F18014_31"

CDS

88..390
 /gene="Atlg19510/F18014_31"
 /codon_start=1
 /product="putative myb-related protein"
 /protein_id="BAC43610.1"
 /db_xref="GI:26453068"
 /translation="MASSSMSSSSWTSKONKMPERALAVYDKTDPDRWQNVAKAVGS
 KSAEVKSHYDLVLEDLMNIEQDLVPLPKYKTVGVSKSRGIDDFDLRLMKMNRQ"

ORIGIN

Query Match 39.7%; Score 151.4; DB 15; Length 580;
 Best Local Similarity 73.2%; Pred. No. 5.5e-27;
 Matches 210; Conservative 0; Mismatches 71; Indels 6; Gaps 1;

```

QY  1  ATGGCGTCAAACTCAAGAAGTTCAATCTCACCATGGACGCTTTAGTCAAAACAAGATGTTTC 60
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB  88  ATGGCCTCTAGTTCTATGAGCTCGAGCTTTCTTGGAGCTCTAAGCAAAACAAGATGTTTC 147

QY  61  GAGAGGCGCTTGGCGAGTTTACGACAAAGGACACACCCGACCGATGGCACAAATGTGCAAAA 120
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB  148  GAGAGGCGCTTGGCGAGTTTACGACAAAGGACACACCCGACCGATGGCACAAATGTGCTAAA 207

QY  121  GCTGTCGAGGAGAACTGTAGAAAGTGAAGCCCACTATGACATTCCTCGTCAGGAT 180
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB  208  GCAGTTGGAAGTAAATCTGACAGGAAGTTAAACGTCACGACATCCTCGTTGAAGAT 267

QY  181  CTCATCAACATCGAGACGTGTCGTCCTTTGGCCCAATTACAAGACCTTCGA-----A 234
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB  268  CTCATGAACATCGAACAGACTTAGTACCTTTGGCTTAATACAAAACCGTCGATTTGA 327

QY  235  TCTAACTCAAGAAGCATCAATGACTTTGACACAAGGTATATACTAA 281
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB  328  AGTAAATCTAGAGGCATCGATGATTTGATTTGAGGTTAATGAAGAA 374

```

RESULT 7

AC025808/c
 LOCUS
 DEFINITION
 Arabidopsis thaliana thaliana BAC F18014 from
 chromosome I, complete sequence.

ACCESSION
 AC025808

VERSION
 AC025808.8 GI:7636235

KEYWORDS
 HTG.

SOURCE
 Arabidopsis thaliana (thale cress)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

SOURCE
ORGANISM

Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE

1 (bases 1 to 506)
Cheuk, R., Chen, H., Kim, C.J., Shinn, P. and Ecker, J.R.

TITLE

Arabidopsis cDNA clones

REFERENCE

2 (bases 1 to 506)
Cheuk, R., Chen, H., Kim, C.J., Shinn, P. and Ecker, J.R.

AUTHORS

Direct Submission

TITLE

Submitted (29-NOV-2003) Salk Institute Genomic Analysis Laboratory

JOURNAL

(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA

FEATURES

Location/Qualifiers

source

1..506

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="4"

/clone="S62918"

/ecotype="Columbia"

/note="This clone is in pUNI 51"

34..279

/note="unknown protein"

/codon_start=1

/product="At4G36570"

/protein_id="AA24137.1"

/db_xref="GI:38566494"

/translation="MASNSMSSASWTRKKNLFRALATYDQTPDRWNVAVGG
KSAEVRHYELLIRVDNIESGRYPHENYRNGNH"

ORIGIN

Query Match 28.5%; Score 108.4; DB 15; Length 506;
Best Local Similarity 68.0%; Pred. No. 2.8e-16;
Matches 151; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 1 ATGGCGTCAACTCAAGAGTTCATCTCACATCGACGTTAGTCAAAACAGATGTC 60

Db 34 ATGGCTTCAACTCAATGAGCTCTAGCGCTCTTGGACACGTAAGGAGAACAAATATT 93

QY 61 GAGAGGGCTTGGCAGTTTACGACAGGACACACCGCCGATGGCACAAATGTGCAAAA 120

Db 94 GAAAGGGCTTGGCTACATATGACCAGACACTCTTGACGTTGCGATACGTTCAAGA 153

QY 121 GCTGTCGAGGAAACCTGTAGAGAGTGAAGGCCCACTATGACATTTCTCGTCAGGAT 180

Db 154 GCCGTGTGGCGGCAATCAGCTGAGAGAGTAAGGCGACACTAGAGCTCCTCATTAGGAT 213

QY 181 CTCATCACATCGACACTGGCTGCTGCTCTTGGCCCAATTAC 222

Db 214 GTCAATGACATGTAGTCAGGCGGTTATCCACATCCCAATTAC 255

RESULT 12

AP008111

LOCUS

AP008111 100108 bp DNA linear HTG 28-DEC-2004

Lotus corniculatus var. japonicus clone Lj705D15, *** SEQUENCING IN

PROGRESS ***, 20 unordered pieces.

ACCESSION

AP008111

VERSION

AP008111.1 GI:56806418

KEYWORDS

HTG; HTGS_PHASE1.

SOURCE

Lotus corniculatus var. japonicus (Lotus japonicus)

ORGANISM

Lotus corniculatus var. japonicus

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
Lotus.

REFERENCE

1
Kaneko, T., Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
Structural Analysis of a Lotus japonicus Genome. XI. Sequence

JOURNAL
REFERENCE

Unpublished
2 (bases 1 to 100108)

AUTHORS

Sato, S.

TITLE

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

Features and Mapping of Nine hundred twenty-one TAC Clones

Unpublished

2 (bases 1 to 100108)

Sato, S.

Direct Submission

Submitted (26-OCT-2004) Shusei Sato, Kazusa DNA Research Institute,

Department of Plant Gene Research; 2-6-7 Kazusa-kamatari, Kisearazu,

Chiba, 292-0818, Japan (E-mail:ssato@kazusa.or.jp,

URL:http://www.kazusa.or.jp/, Tel:81-438-52-3935(ex.2337),

Fax:81-438-52-3934)

* NOTE: This is a 'working draft' sequence. It currently

* consists of 20 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence.

* as soon as it is available and the accession number will

* be preserved.

* 1 5202: contig of 5202 bp in length

* 5203: gap of unknown length

* 5303: contig of 1264 bp in length

* 5567: gap of unknown length

* 6667: contig of 1094 bp in length

* 7761: gap of unknown length

* 7861: contig of 764 bp in length

* 8625: gap of unknown length

* 8725: contig of 883 bp in length

* 9608: gap of unknown length

* 9708: contig of 1834 bp in length

* 11542: gap of unknown length

* 11642: contig of 2154 bp in length

* 13796: gap of unknown length

* 13896: contig of 1847 bp in length

* 15743: gap of unknown length

* 15843: contig of 3189 bp in length

* 19032: gap of unknown length

* 19132: contig of 2679 bp in length

* 21811: gap of unknown length

* 21911: contig of 2757 bp in length

* 24668: gap of unknown length

* 24768: contig of 5539 bp in length

* 30307: gap of unknown length

* 30407: contig of 6179 bp in length

* 36586: gap of unknown length

* 36686: contig of 5704 bp in length

* 42390: gap of unknown length

* 42490: contig of 8402 bp in length

* 50891: gap of unknown length

* 50992: contig of 9150 bp in length

* 60142: gap of unknown length

* 60242: contig of 8837 bp in length

* 69079: gap of unknown length

* 69179: contig of 10310 bp in length

* 79489: gap of unknown length

* 79589: contig of 18889 bp in length

* 98478: gap of unknown length

* 98578: contig of 1531 bp in length.

Location/Qualifiers

1..100108

/organism="Lotus corniculatus var. japonicus"

/mol_type="genomic DNA"

/var_tety="japonicus"

/db_xref="taxon:34305"

/clone="Lj705D15"

/note="TAC clone:TM1490, synonym:Lotus japonicus"

5203..5302

/estimated_length=unknown

6567..6666

/estimated_length=unknown

7761..7860

/estimated_length=unknown

8625..8724

FEATURES

source

gap

gap

gap

gap

```
/estimated_length=unknown
9608..9707
/estimated_length=unknown
11542..11541
/estimated_length=unknown
13796..13895
/estimated_length=unknown
15743..15842
/estimated_length=unknown
19032..19131
/estimated_length=unknown
21811..21910
/estimated_length=unknown
24668..24767
/estimated_length=unknown
30307..30406
/estimated_length=unknown
36586..36685
/estimated_length=unknown
42390..42489
/estimated_length=unknown
50892..50991
/estimated_length=unknown
60142..60241
/estimated_length=unknown
69079..69178
/estimated_length=unknown
79489..79588
/estimated_length=unknown
98478..98577
/estimated_length=unknown

ORIGIN
Query Match      28.5%; Score 108.4; DB 14; Length 100108;
Best Local Similarity 58.5%; Pred. No. 2.8e-16;
Matches 207; Conservative 0; Mismatches 146; Indels 1; Gaps 1;

QY 1 ATGGCGTCAAACTCAGAAGTTCAATCTCACCATGGAGCTTTAGTCAAAACAAGATGTTTC 60
Db 74473 ATGGCATCAAGCTCAATGTCATCTATGCTCATGGAATGCTAAGGACAATAAGCCGTTT 74532
QY 61 GAGAGGCGCTTGGCAGTTTACGACAGGACACACCCGACCGATGGCAATGCGCAAAA 120
Db 74533 GAAAGGGCTCTAGCTGTTTATGACAGGNAACCCCGACAGATGGTACAAATGTGCTCAA 74592
QY 121 GCTGTGCGAGGAAACTGTAGAGAAAGTGAAGCGCCACTATGACATTTCTCGTCGAGGAT 180
Db 74593 GCTGTTGGTGGGAAGACCCAGAGGAAGTGAAGGGCACTATGAGCTTCTCCTTAGGGAT 74652
QY 181 CTCATCAACATCGAGACTGTCGTCGTCCTTTGCCCAATTACAGACCTTCGAATCTTAAC 240
Db 74653 GTTAGGCACATTTGAATCAGGCGCATGTTCCATTTCCAAACTACAAGAACACTGGA-GGGTC 74711
QY 241 TCAAGAGCATCAATGACTTTTGACACAGGATATATACTAAATATCTATATATGATGCTC 300
Db 74712 AGAACAGGAAAAGGTAGTGACACATAAAACTGTATCTTCATATTCCTCACTTGAACAAA 74771
QY 301 TCGATATATTTTGATAATCATCTTCAGTGATTTTGAGAAATTTCTCTCAAAAAGTT 354
Db 74772 TGCMAATATGCAATCTTCACTTCTGCGCTGTGCTAAGTTGCTATATAGTT 74825

RESULT 13
AP004546/c
LOCUS
DEFINITION
Lotus corniculatus var. japonicus genomic DNA, chromosome 1,
clone:Lj743N05, TM0178, complete sequence.
ACCESSION
AP004546
VERSION
AP004546.1 GI:17736913
KEYWORDS
HTG.
SOURCE
Lotus corniculatus var. japonicus (Lotus japonicus)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
Lotus.
1
Sato, S., Kaneko, T., Nakamura, Y., Asamizu, E., Kato, T. and Tabata, S.
Structural Analysis of a Lotus japonicus Genome. I. Sequence
Features and Mapping of Fifty-six TAC clones which cover the 5.4 Mb
Regions of the Genome
Unpublished
2 (bases 1 to 97268)
Nakamura, Y.
Direct Submission
Submitted (13-DEC-2001) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research, 1532-3, Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail:ynakam@kazusa.or.jp,
URL:http://www.kazusa.or.jp, Tel:81-438-52-3935,
Fax:81-438-52-3934)
Location/Qualifiers
1..97268
/organism="Lotus corniculatus var. japonicus"
/mol_type="genomic DNA"
/variety="japonicus"
/db_xref="taxon:34305"
/chromosome="1"
/clone="Lj743N05"
/clone_lib="LjT library"
/note="TAC clone:TM0178"
synonym: Lotus japonicus"

ORIGIN
Query Match      28.0%; Score 106.6; DB 15; Length 97268;
Best Local Similarity 70.6%; Pred. No. 7.9e-16;
Matches 142; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 28 TCACCATGGACGTTTGTAGTCAAAACAAGATGTTTCGAGAGGCGCTTGGCAGTTTACGACAAG 87
Db 32507 TCTACTTGGAGTCCAAAACAGACAAAGTTGTTTGAGAGGACCTTCGAAAATATGACGAG 32448
QY 88 GACACACCCGACGATGGCACAATGTGGCAAAAGCTGTCGAGGGGAAAACCTGTAGAGAA 147
Db 32447 GACACCCCTGAGCGGTGGCAAAATGTCCAAAGCAGTTGGTGGCAAAATCAGTGGAGGAA 32388
QY 148 GTGACGCGCACTATGACATTTCTCGTCGAGGATCTCATCAACATCAGACTGCTCGTGTTC 207
Db 32387 GTCAAGAGACACTATGAGAGACTTTGGAGGATCTCAAAACACATTCGATCTGCCCATGTT 32328
QY 208 CCTTTGCCCAATTACAAGACC 228
Db 32327 CCTTTCCCACTACAAATCC 32307

RESULT 14
AP008039/c
LOCUS
DEFINITION
Lotus corniculatus var. japonicus chromosome 1 clone LjT10P24, ***
SEQUENCING IN PROGRESS ***, 24 unordered pieces.
ACCESSION
AP008039
VERSION
AP008039.1 GI:56806347
KEYWORDS
HTG; HTGS, PHASE1.
SOURCE
Lotus corniculatus var. japonicus (Lotus japonicus)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
Lotus.
1
Kaneko, T., Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
Structural Analysis of a Lotus japonicus Genome. XI. Sequence
Features and Mapping of Nine hundred twenty-one TAC Clones
Unpublished
2 (bases 1 to 102550)
Sato, S.
Direct Submission
```

JOURNAL

Submitted (26-OCT-2004) Shusei Sato, Kazusa DNA Research Institute,
Department of Plant Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu,
Chiba, 292-0818, Japan (E-mail:ssato@kazusa.or.jp,
URL:http://www.kazusa.or.jp/, Tel:81-438-52-3935(ex.2337),
Fax:81-438-52-3934)

COMMENT

* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 3089: contig of 3089 bp in length
* 3090 3189: gap of unknown length
* 3190 4482: contig of 1293 bp in length
* 4483 4582: gap of unknown length
* 4583 5591: contig of 1009 bp in length
* 5592 5691: gap of unknown length
* 5692 6980: contig of 1289 bp in length
* 6981 7080: gap of unknown length
* 7081 8737: contig of 1657 bp in length
* 8738 8837: gap of unknown length
* 8838 10414: contig of 1577 bp in length
* 10415 10514: gap of unknown length
* 10515 11793: contig of 1279 bp in length
* 11794 11893: gap of unknown length
* 11894 14747: contig of 2854 bp in length
* 14748 14847: gap of unknown length
* 14848 17118: contig of 2271 bp in length
* 17119 17218: gap of unknown length
* 17219 19795: contig of 2577 bp in length
* 19796 19895: gap of unknown length
* 19896 23048: contig of 3153 bp in length
* 23049 23149: gap of unknown length
* 23150 25125: contig of 1976 bp in length
* 25126 25157: gap of 33 bp
* 25158 25546: contig of 389 bp in length
* 25547 25646: gap of unknown length
* 25647 28751: contig of 3105 bp in length
* 28752 28851: gap of unknown length
* 28852 32078: contig of 3227 bp in length
* 32079 32178: gap of unknown length
* 32179 37013: contig of 4835 bp in length
* 37014 37113: gap of unknown length
* 37114 41223: contig of 4110 bp in length
* 41224 41324: gap of unknown length
* 41325 46917: contig of 5494 bp in length
* 46918 46917: gap of unknown length
* 46919 54494: contig of 7577 bp in length
* 54495 54594: gap of unknown length
* 54595 64935: contig of 10341 bp in length
* 64936 65035: gap of unknown length
* 65036 74018: contig of 8983 bp in length
* 74019 74118: gap of unknown length
* 74119 88264: contig of 14146 bp in length
* 88265 88364: gap of unknown length
* 88365 100720: contig of 12356 bp in length
* 100721 100820: gap of unknown length
* 100821 102550: contig of 1730 bp in length.
Location/Qualifiers
1..102550
/organism="Lotus corniculatus var. japonicus"
/mol_type="genomic DNA"
/variety="japonicus"
/db_xref="taxon:34305"
/chromosome="1"
/clone="LjT10P24"
/clone_lib="LjT library"
/note="FAC clone: FW1388, synonym: Lotus japonicus"
3090..3189
/estimated_length=unknown
4483..4582

FEATURES
source

1..102550

/organism="Lotus corniculatus var. japonicus"

/mol_type="genomic DNA"

/variety="japonicus"

/db_xref="taxon:34305"

/chromosome="1"

/clone="LjT10P24"

/clone_lib="LjT library"

/note="FAC clone: FW1388, synonym: Lotus japonicus"

3090..3189

/estimated_length=unknown

4483..4582

gap

gap

gap /estimated_length=unknown
5592..5691
/estimated_length=unknown
5981..7080
/estimated_length=unknown
8738..8837
/estimated_length=unknown
10415..10514
/estimated_length=unknown
11794..11893
/estimated_length=unknown
14748..14847
/estimated_length=unknown
17119..17218
/estimated_length=unknown
19796..19895
/estimated_length=unknown
23049..23148
/estimated_length=unknown
25125..25157
/estimated_length=33
25547..25646
/estimated_length=unknown
28752..28851
/estimated_length=unknown
32079..32178
/estimated_length=unknown
37014..37113
/estimated_length=unknown
41224..41323
/estimated_length=unknown
46818..46917
/estimated_length=unknown
54495..54594
/estimated_length=unknown
64936..65035
/estimated_length=unknown
74019..74118
/estimated_length=unknown
88265..88364
/estimated_length=unknown
100721..100820
/estimated_length=unknown
ORIGIN

Query Match 28.0%; Score 106.6; DB 14; Length 102550;
Best Local Similarity 70.6%; Pred. No. 7.9e-16;
Matches 142; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 28 TCACCATGGACGTTTAGTCAAAACAAGATGTCGAGAGGGCGCTTCGACAG 87
DB 96874 TCTACTTGGAGTCCAAACAAGACAAAGTGTGAGAGAGACATTCGACAG 96815
QY 88 GACACACCCGACCGATGGCAAAATGTGCAAAAGCTGCGAGGGGAAACTGTAGAAGAA 147
DB 96814 GACACCCCTGAGCGGTGGCAAAATGTGTCCAAGACAGTTGGTGGCAATCATGTGAGGAA 96755
QY 148 GTGAAGCGCCACTATGACATTCCTCGTCGAGGATCTCAACATCATGAGACTGCTGTC 207
DB 96754 GTCAAGAGACACTATGAGAGACTCTTGGAGGATCTCAAAACACATTCGAGTCTGCGCATGTT 96695
QY 208 CCTTTGCCCAATTACAGACC 228
DB 96694 CCCTTTCCCACTACAAATCC 96674

RESULT 15

AP007808/c

LOCUS

DEFINITION

AP007808

AP007808

AP007808

AP007808

AP007808 108878 bp DNA linear HTG 28-DEC-2004
Lotus corniculatus var. japonicus chromosome 1 clone LjB04J07, ***
SEQUENCING IN PROGRESS ***, 24 unordered pieces.
AP007808
AP007808.1 GI:56806118

KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

HTG; HTGS_PHASE1.
Lotus corniculatus var. japonicus (Lotus japonicus)
Lotus corniculatus var. japonicus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
Rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
Lotus.

1
Kaneko, T., Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
Structural Analysis of a Lotus japonicus Genome. XI. Sequence
Features and Mapping of Nine hundred twenty-one TAC Clones
Unpublished
2 (bases 1 to 108878)
Sato, S.
Direct Submission
Submitted (26-OCT-2004) Shusei Sato, Kazusa DNA Research Institute,
Department of Plant Gene Research: 2-6-7 Kazusa-kamatari, Kisarazu,
Chiba, 292-0818, Japan (E-mail: ssato@kazusa.or.jp, Kisarazu,
URL: http://www.kazusa.or.jp/, Tel: 81-438-52-3935(ex.2337),
Fax: 81-438-52-3934)

COMMENT
* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 4606: contig of 4606 bp in length
* 4607 4706: gap of unknown length
* 4707 5430: contig of 744 bp in length
* 5451 5550: gap of unknown length
* 5551 6707: contig of 1157 bp in length
* 6708 6807: gap of unknown length
* 6808 8058: contig of 1251 bp in length
* 8059 8158: gap of unknown length
* 8159 9271: contig of 1113 bp in length
* 9272 9372: gap of unknown length
* 9372 10504: contig of 1132 bp in length
* 10504 12079: gap of unknown length
* 12080 12179: contig of 1476 bp in length
* 12180 13219: gap of unknown length
* 13220 15621: contig of 2302 bp in length
* 15622 15721: gap of unknown length
* 15722 17539: contig of 1818 bp in length
* 17540 17639: gap of unknown length
* 17640 19483: contig of 1843 bp in length
* 19483 19583: gap of unknown length
* 19583 20911: contig of 1328 bp in length
* 20911 23211: gap of unknown length
* 23211 23312: contig of 2201 bp in length
* 23312 26479: contig of 3168 bp in length
* 26480 26579: gap of unknown length
* 26580 29861: contig of 3281 bp in length
* 29861 36032: gap of unknown length
* 36032 36132: contig of 6072 bp in length
* 36132 36333: gap of unknown length
* 36333 39621: contig of 3488 bp in length
* 39621 45217: gap of unknown length
* 45217 45318: contig of 5497 bp in length
* 45318 49895: gap of unknown length
* 49895 58773: contig of 4578 bp in length
* 58773 58874: gap of unknown length
* 58874 67741: contig of 8867 bp in length
* 67741 79594: gap of unknown length
* 79594 93084: contig of 11754 bp in length
* 93084 93084: contig of 13390 bp in length
* 79695 79695

FEATURES
source

* 93085 93184: gap of unknown length
* 93185 108878: contig of 15694 bp in length.
1. 108878
Location/Qualifiers
/organism="Lotus corniculatus var. japonicus"
/mol_type="genomic DNA"
/variant="japonicus"
/db_xref="taxon:34305"
/chromosome="1"
/clone_lib="LjB library"
/note="BAC clone:BM1105, synonym:Lotus japonicus"
4607..4706
/estimated_length=unknown
5451..5550
/estimated_length=unknown
6708..6807
/estimated_length=unknown
8059..8158
/estimated_length=unknown
9272..9371
/estimated_length=unknown
10504..10603
/estimated_length=unknown
12080..12179
/estimated_length=unknown
13220..13319
/estimated_length=unknown
15622..15721
/estimated_length=unknown
17540..17639
/estimated_length=unknown
19483..19582
/estimated_length=unknown
20911..21010
/estimated_length=unknown
23212..23311
/estimated_length=unknown
26480..26579
/estimated_length=unknown
29861..29960
/estimated_length=unknown
36033..36132
/estimated_length=unknown
39621..39720
/estimated_length=unknown
45218..45317
/estimated_length=unknown
49896..49995
/estimated_length=unknown
58774..58873
/estimated_length=unknown
67741..67840
/estimated_length=unknown
79595..79694
/estimated_length=unknown
93085..93184
/estimated_length=unknown

ORIGIN

Query Match 28.0%; Score 106.6; DB 14; Length 108878;
Best Local Similarity 70.6%; Pred.No. 7.9e-16;
Matches 142; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
QY 28 TCACCATGGACGTTTGTAGTCAAAACAAGATGTTTCGAGAGGCGCTTGGCAGTTTACGACAAG 87
58515 TCTACTTGGAGTCCAAAACAGAACAAAGTTGTTTGAGAGAGCACTTGCAAAATATGACGAG 58456
Db
QY 88 GACACACCGACCGATGGCACAATGTGGCAAAAGCTGTCCGAGGGGAAAAGTGTAGAAGAA 147
Db 58455 GACACCCCTGAGCGGTGGCAAAATGTGTCCAAAGACAGTTGGTGGCAATCAGTGGAGGAA 58396
QY 148 GTGAAGCGCCACTATGACATTCTCGTCGAGGATCTCATCAACATCGAGACTGGTCTGTGTC 207

Db 58395 GTCAAGAGACACTATGAGAGACTCTTTGGAGGATCTCAACACACATTGAGTCTGCGCATGTT 58336
Qy 208 CCTTGGCCCAATTACAGACC 228
Db 58335 CCTTTCCCACTACAAATCC 58315

Search completed: March 10, 2006, 19:26:15
Job time : 2280 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2006, 18:36:58 ; Search time 136 Seconds
(without alignments)
4979.787 Million cell updates/sec

Title: US-10-697-787-1
Perfect score: 381
Sequence: 1 atggcgctcaactcaagaag.....gttatattcttggtttaa 381

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues
Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/1_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/H_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq: *
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq: *
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq: *
9: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	69.6	18.3	574	US-09-640-211A-1534	Sequence 1534, Ap
2	58.4	15.3	357	US-09-640-211A-1581	Sequence 1581, Ap
3	50.4	13.2	383	US-09-640-211A-1310	Sequence 1310, Ap
4	40.2	10.6	100863	US-09-949-016-17031	Sequence 17031, A
C 5	38.8	10.2	399	US-09-621-976-8976	Sequence 8976, Ap
6	37.4	9.8	89892	US-09-949-016-13667	Sequence 13667, A
C 7	37.4	9.8	670689	US-09-949-016-12505	Sequence 12505, A
C 8	37.4	9.8	670690	US-09-949-016-14207	Sequence 14207, A
C 9	37	9.7	1141	US-09-806-708B-22	Sequence 22, Appl
10	36.8	9.7	1689	US-08-447-500-11	Sequence 11, Appl
11	36.8	9.7	1689	US-08-453-866-11	Sequence 11, Appl
12	36.8	9.7	2385	US-08-454-097-9	Sequence 9, Appl
13	36.8	9.7	2385	US-08-468-036-41	Sequence 41, Appl
14	36.8	9.7	2385	US-08-376-843-41	Sequence 41, Appl
15	36.8	9.7	2385	US-08-185-359-9	Sequence 9, Appl
16	36.8	9.7	58356	US-09-949-016-15563	Sequence 15563, A
17	36.6	9.6	601	US-09-949-016-103973	Sequence 103973, A
18	36.6	9.6	8868	US-08-815-809-7	Sequence 7, Appl
C 19	36.6	9.6	9870	US-09-245-928A-15	Sequence 15, Appl
20	36.6	9.6	19877	US-08-816-155B-8	Sequence 8, Appl
21	36.6	9.6	19877	US-09-073-587-8	Sequence 8, Appl
22	36.6	9.6	84839	US-09-949-016-15816	Sequence 15816, A
23	36.6	9.6	147382	US-09-949-016-14624	Sequence 14624, A
C 24	36.4	9.6	16573	US-09-949-016-14876	Sequence 14876, A

C 25	36	9.4	601	3	US-09-949-016-78876	Sequence 78876, A
C 26	36	9.4	784019	3	US-09-949-016-14033	Sequence 14033, A
C 27	36	9.4	828152	3	US-09-949-016-12777	Sequence 12777, A
C 28	35.6	9.3	429	3	US-09-621-976-9543	Sequence 9543, Ap
29	35.6	9.3	1860	3	US-09-919-497-50	Sequence 50, Appl
30	35.6	9.3	2069	3	US-09-949-016-5801	Sequence 5801, Ap
31	35.4	9.3	332	3	US-09-640-211A-1824	Sequence 1824, Ap
32	35.2	9.2	32392	3	US-09-662-254B-27	Sequence 27, Appl
C 33	34.8	9.1	1230	3	US-09-140-466-1	Sequence 1, Appl
34	34.6	9.1	1784	3	US-09-486-581-1	Sequence 1, Appl
35	34.6	9.1	1784	3	US-10-213-797-1	Sequence 1, Appl
36	34.6	9.1	38346	3	US-09-949-016-14502	Sequence 14502, A
37	34.6	9.1	39318	3	US-09-949-016-13798	Sequence 13798, A
C 38	34.6	9.1	43117	3	US-09-949-016-17589	Sequence 17589, A
39	34.6	9.1	52199	3	US-09-949-016-12763	Sequence 12763, A
40	34.6	9.1	52203	3	US-09-949-016-16330	Sequence 16330, A
41	34.4	9.0	290	3	US-10-131-827-8636	Sequence 8636, Ap
42	34.4	9.0	601	3	US-09-949-002-8929	Sequence 8929, Ap
C 43	34.4	9.0	1885	3	US-09-662-254B-41	Sequence 41, Appl
44	34.4	9.0	50000	3	US-09-662-254B-24	Sequence 24, Appl
C 45	34.4	9.0	80411	3	US-09-949-016-15777	Sequence 15777, A

ALIGNMENTS

RESULT 1
US-09-640-211A-1534/c
; Sequence 1534, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1534
; LENGTH: 574
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-1534

Query Match	18.3%	Score 69.6;	DB 3;	Length 574;
Best Local Similarity	59.7%	Pred. No. 2.3e-10;		
Matches 117;	Conservative 0;	Mismatches 79;	Indels 0;	Gaps 0;
QY 28	TCACCATGGAGCTTTAGTCAAAACAAAGATGTTGAGAGGGCTTGGCAGATTTACGACAAG 87			
Db 431	TCCTCGTGGAGTAAGGAACAGGACGCTTCGAGAATGCCCTGGCTACCCACATTCGAG 372			
QY 88	GACACACCCGACCGATGCCAATGTGCAAAAGCTGCGGAGGGAACACTGTAGAAGAA 147			
Db 371	GTTCCTCCCGAGTTGTTGGAGAAAGTAGCTGCCGACGTGCCCGGCAAAACCTTAGAGGAG 312			
QY 148	GTGAAGCGCCACTATGACATCTCTGTCGAGGATCTCATCAACATCGAGACTGGTGGTGTGTC 207			
Db 311	ATTAAGCTCCACTACGAGGACCTGGTGAAGATGTCATAGATAGAGGCTGGCCGCGTG 252			
QY 208	CCTTGGCCCAATTACA 223			
Db 251	CCTCTGCGCTCTTACA 236			

RESULT 2
US-09-640-211A-1581
; Sequence 1581, Application US/09640211A
; Patent No. 6833446

Db 277 RMYRRAMRSGKSGGSGYRMYAGYRSSRWYSAMWRKQKQKQKGRSSWGRSTGY 218
QY 109 AATGTGCAAAAGCTGCGAGGAGAACTGTAGAAGAGTGAAGCGCCACTATGACATT 168
Db 217 YAMWYKSWCTSRWYKWKRRKCTSTKRTCYRSTYKWKAYYTKRRKWTWT 158
QY 169 CTCGTGAGGATCTCATCAACATCGAGACTGGTGGTCCCTTTGCCCAATTACAAGACC 228
Db 157 YYYKYSMSKKTWRMKTAYYTKRWK-MTRTKWTCTWCKCTTYMAGTMMYRYRYW 99
QY 229 TTGCAATCTCAACAAGACATCAATGACTTTGCACAAAGGTATATATACTAAATCTA 288
Db 98 YYAKRAKWSKRCWTSTTCYCMKMYMAKKCWSYMWMSNMKMGKSMWKKWTYYYYYMMKWS 39
QY 289 TATATGATGCTCTCGATATATATTGTATAAT 318
Db 38 KMTYNSMWSYARKCWRTYAKTYTWTCTMT 9

RESULT 6
US-09-949-016-13667
; Sequence 13667, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13667
; LENGTH: 89892
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(89892)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13667

Query Match
Best Local Similarity 9.8%; Score 37.4; DB 3; Length 89892;
Matches 80; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
QY 224 AGACCTTCGAATCTCACTCAAGACATCAATGACTTTGCACAAAGGTATATACTAAAT 283
Db 66602 AGATATGCACTCTCTATATAGACATAGATATAAAATATATGACATATTTAAATAAT 66661
QY 284 ATCTATATATGATGCTCTCGATATATTTTGAATCAATCTAGTATTTTGAGAAATCT 343
Db 66662 TTATATATATGATGATACATATATTAATTTACATATATTAATATATATATACATTAT 66721
QY 344 CTCAAAAAGTCTTGTGAAGTTATATTTCTTT 374
Db 66722 TTTAATAAATTATTATTATTATGAGCCTT 66752

RESULT 7
US-09-949-016-12505/c
; Sequence 12505, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12505
; LENGTH: 670689
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(670689)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12505
Query Match
Best Local Similarity 9.8%; Score 37.4; DB 3; Length 670689;
Matches 80; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
QY 220 TACAAGACCTTCGAATCTCACTCAAGACATCAATGACTTTGCACAAAGGTATATACT 279
Db 544773 TAATAGACATTTTAAGCAACAAGAAATTAATTTGTTACCTTTAAATATGATAATGAAT 544714
QY 280 AAATATCTATATATGATGCTCTCGATATATTTTGAATCAATCTAGTATTTTGAGAAA 339
Db 544713 GAAGCTCTATATGCCAATTTGATGATATATATAGCATTCCCATTAAGCACATTTCTCGATA 544654
QY 340 TTCTCTCAAAAAGTCTTGTGAAGTTATATTT 370
Db 544653 TATTTCTCAAAAAGTCTTGTGAAGTT 544623

RESULT 8
US-09-949-016-14207/c
; Sequence 14207, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14207
; LENGTH: 670690
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(670690)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14207

Query Match
Best Local Similarity 9.8%; Score 37.4; DB 3; Length 670690;
Matches 80; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
QY 220 TACAAGACCTTCGAATCTCACTCAAGACATCAATGACTTTGCACAAAGGTATATACT 279

Db 544773 TAAATAGACATTTTAAAGCACAGCAAAATTAATATGTACCTTTAAATATGATGAATGAAC 544714
Qy 280 AAATATCTATATATGATCTCTCGATATATTTTGATAATCTTCTAGTGTAGTTTGAGAAA 339
Db 544713 GAAGCTCTATATGCCAATGTTATGATATATACACATTCACATTTCCCATTAAGCACATTTCTCGATA 544654
Qy 340 TTCTCTCAAAAGTCTTCTGTAAGTTATATTT 370
Db 544653 TATCTGCAAAAAGTCTTGTGTTCTTAAAGTT 544623

RESULT 9
US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. PAF1 promoters
US-09-806-708B-22

Query Match 9.7%; Score 37; DB 3; Length 1141;
Best Local Similarity 14.7%; Pred. No. 0.84;
Matches 44; Conservative 91; Mismatches 164; Indels 0; Gaps 0;
Qy 71 TGGCAGTTTACGACAAAGGACACACCCGACGCGCAATGTGGCAAAAGCTGTCGGAG 130
Db 956 TMBGWADTAGKCNNNNNWTTDVRMAKAKNNNNNAYWTACYNRAATNNKMAHWM 897
Qy 131 GGAATACTGTAGAGAGAGTGAAGCGCACATATGACATCTCGTCGAGGATCTCATCAACA 190
Db 896 KWTGHAHSKRTRHRTTCRTKYNNNNNNARTVYVYHHAARRMNAWWTCTNNNNNNNN 837
Qy 191 TCGAGACTGTCGTGTCCTTCCCAATACAGACCTTCGAATCTCAAGACCTCAAGACGA 250
Db 836 NNACRNRTRTWABWKHSWNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 777
Qy 251 TCAATGACTTTGACACAAAGGTATATACTAAATATCTATATATGATGCTCTCGATATAT 310
Db 776 TCNNYMAAATTTTDCYCTWMTWTDMMTWBTTTNNNNNNNNNNNNNNNNNNNNNN 717
Qy 311 TTGATAATCTTCTAGTGAATTTTGAGAAATCTCTCAAAAAGTTCTTGTAGGTATAT 369
Db 716 NNNNNNNWKAYYAHATNNWGCWNNNTDARRTNNNTVMRRRWMNTKTRWYSTRHHYT 658

RESULT 10
US-08-447-500-11
; Sequence 11, Application US/08447500
; Patent No. 5627064
; GENERAL INFORMATION:
; APPLICANT: Hoechst, Merl F.
; TITLE OF INVENTION: PROTEIN KINASES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East, Suite 500
; CITY: Los Angeles
; STATE: CA

; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,500
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/008,001
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell Ph.D., John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-2458
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1689 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: Protein Kinase
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 294..1385
US-08-447-500-11

Query Match 9.7%; Score 36.8; DB 2; Length 1689;
Best Local Similarity 56.7%; Pred. No. 1.1;
Matches 68; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
Qy 246 AAGCATCAATGACTTTGACACAAAGGTATATACTAAATATCTATATATGATGCTCTCGAT 305
Db 1551 AAACCCAAATGCTTTTCATATATATGATATGATATGATATATATACATATATGCTGT 1610
Qy 306 ATATTTTGATATCATCTTAGTGATTTTGAGAAATCTCTCAAAAAGTTCTTCTTAAGTTA 365
Db 1611 ATATTTATATCATCTTCTTGGGATTTGGGTCATTTTTTTAAACAACCTGATCTTTTTA 1670

RESULT 11
US-08-453-866-11
; Sequence 11, Application US/08453866
; Patent No. 5756289
; GENERAL INFORMATION:
; APPLICANT: Hoechst, Merl F.
; TITLE OF INVENTION: PROTEIN KINASES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East, Suite 500
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,866
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

; ; TELECOMMUNICATION INFORMATION:

US-08-468-036-41

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model.

Run on: March 10, 2006, 18:32:51 ; Search time 2647 Seconds
(without alignments)
6734.370 Million cell updates/sec

Title: US-10-697-787-1

Perfect score: 381

Sequence: 1 atggcgctcaactcaagaag.....gttatattcttgggttaa 381

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_est3:*
- 4: gb_est4:*
- 5: gb_est5:*
- 6: gb_est6:*
- 7: gb_est7:*
- 8: gb_est8:*
- 9: gb_est9:*
- 10: gb_est10:*
- 11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	181	47.5	256	9	CC179487 SALK_0699
2	146.8	38.5	793	9	BH947619 obu80f11
3	137.2	36.0	446	8	L38243 BNAF0581E M
4	128.8	33.8	456	7	CO052301 Mdfw20570
5	124.2	32.6	632	6	CF230621 PtaC0010H
6	122.6	32.2	662	7	CV277080 WS0142.B2
7	121.8	32.0	728	5	BU890694 P040E03 P
8	121.2	31.8	457	6	CD669972 eepic.pk0
9	121	31.8	516	7	CV278823 WS0147.B2
10	121	31.8	588	7	CV283381 WS0187.B2
11	120.4	31.6	624	7	CV232994 WS0199.B2
12	120.4	31.6	627	7	CV230387 WS01916.B
13	119.8	31.4	626	6	CB920052 VMD058F11
14	119.8	31.4	662	6	CB919185 VMD042D03
15	118.2	31.0	474	5	BU668323 MC01026A1
16	117.8	30.9	349	5	BU822353 UB36DPB08
17	117	30.7	513	7	CO997831 pam01-16m
18	115.4	30.3	599	8	CV882244 Mdlv3_403
19	115.4	30.3	653	8	DR997716 Mdfb8004M
20	115.4	30.3	659	8	DR990507 Mdlr7005H
21	115.4	30.3	660	8	DR995864 Mda9011J
22	115.4	30.3	688	7	CV186968 Mdlv2_401

23	113.6	29.8	629	8	DR994824
24	113.4	29.8	577	1	AJ768009
25	112.4	29.5	460	2	BG726181
26	112.2	29.4	497	6	CF603713
c 27	111.4	29.2	476	6	CB035850
c 28	110.6	29.0	611	6	CF373337
29	110	28.9	749	6	CB972859
30	109.8	28.8	492	6	CB002633
31	109.8	28.8	492	6	CB002772
32	109.8	28.8	492	6	CB004008
33	109.8	28.8	509	6	CB344615
34	109.8	28.8	513	6	CB342316
35	109.8	28.8	528	6	CB811699
36	109.8	28.8	585	3	BM437293
37	109.8	28.8	613	6	CB345061
38	109.8	28.8	618	6	CB344529
39	109.8	28.8	623	6	CB343483
c 40	109.8	28.8	636	6	CB343539
41	109.8	28.8	637	6	CB811769
42	109.8	28.8	697	6	CB343846
43	108.8	28.6	556	1	AJ320038
c 44	108.8	28.6	557	1	AJ320048
45	108.8	28.6	558	1	AJ320067

ALIGNMENTS

RESULT 1
CC179487
LOCUS CCI179487 256 bp DNA linear GSS 02-MAY-2003
DEFINITION SALK_069941.39.90.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_069941.39.90.X, genomic
survey sequence.
ACCESSION CCI179487 GI:30318038
VERSION CCI179487
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 256)
AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
Gadriab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
Shinn,P., Zimmerman,J. and Ecker,J.R.
TITLE A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
JOURNAL Unpublished (2001)
COMMENT Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salic.edu
This is single pass sequence recovered from the left border of
TDNA. This sequence lies within 300 bases of the 5' end of
CC179487.
Class: TDNA tagged.
Location/Qualifiers
1. 256
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="SALK_069941.39.90.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can

```

be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN
Query Match          47.5%; Score 181; DB 9; Length 256;
Best Local Similarity 94.9%; Pred. No. 9.7e-38;
Matches 187; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 ATGGCGTCAAACTCAAGAGTTCAATCTCACCATGGACGTTTAGTCAAAACAAGATGTTTC 60
Db 60 ATGGCGTCAAACTCAAAAGTTCAATCTCACCATGGACGTTTAGTCAAAACAAGATGTTTC 119

QY 61 GAGAGGCGCTTGGCAGTTTACGACAGGACACACCCGACCGATGGACATGTCGCAAAA 120
Db 120 GAGAGGCGCTTGGCAGTTTACGACAGGACACACCCGACCGATGGACATGTCGCAAAA 179

QY 121 GCTGTCGGAGGAAAACTGTAGAAGAGTGAAGCGCCACTATGACATTTCTCGTCGAGGAT 180
Db 180 GCTGTCGGAGGAAAACTGTATAGAAGTGAAGCGCCTCTATGACATTTCTCGTCCAAGAA 239

QY 181 CTCATCAACATCGAGAC 197
Db 240 CTCATCAACATCGAGAC 256

RESULT 2
BH947619/c
LOCUS          793 bp DNA linear GSS 01-OCT-2002
DEFINITION    obu80f11.b1 B.oleracea002 Brassica oleracea genomic, genomic survey
               sequence.
ACCESSION     BH947619
VERSION       BH947619.1 GI:23427679
KEYWORDS      GSS.
SOURCE        Brassica oleracea
ORGANISM      Brassica oleracea
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
               rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE     1 (bases 1 to 793)
AUTHORS       Delehaunty,K., Fellw,G., Fulton,L., McCombie,W.R., Miner,T.,
               Nash,W., Rabinowicz,P.D. and Wilson,R.K.
TITLE         Whole genome shotgun reads from Brassica oleracea
JOURNAL       Unpublished (2002)
COMMENT       Contact: Richard K. Wilson
               Genome Sequencing Center
               Washington University School of Medicine
               Email: submissions@watson.wustl.edu
               Plate: obu80 row: f column: 11
               Seq primer: -21UPpOT forward
               Class: shotgun
               High quality sequence start: 112
               High quality sequence stop: 551.
               Location/Qualifiers
                   1..793
                       /organism="Brassica oleracea"
                       /mol_type="genomic DNA"
                       /db_xref="taxon:3712"
                       /clone_lib="B.oleracea002"
                       /notes="Vector: pOTw13; whole genome shotgun library from
               flowering buds. DNA was purified from a crude nuclear
               prep using Brassica oleracea T01000DH3 buds provided by
               Thomas Oborn at the University of Wisconsin. Genomic
               DNA was provided by Pablo Rabinowicz (CSHL) and the
               shotgun library prepared at Washington University Genome
               Sequencing Center."

ORIGIN
Query Match          38.5%; Score 146.8; DB 9; Length 793;
Best Local Similarity 90.2%; Pred. No. 1.8e-28;
Matches 157; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 107 ACAATGGCAAAAGCTGTCGGAGGAAAACTGTAGAAGAGTGAAGCGCCACTATGACA 166
Db 545 AAATCTCTCAACGCTGTCGGAGGAAATCTGCAGACGAAGCGTCACTATGACA 486

```

```

QY 167 TTCTCGTCGAGGATCTCATCAATCGAGACTGGTGGTCCCTTTGCCCAATTACAAGA 226
Db 485 TTCTCGTCGAGGATCTCATCAATCGAAATCGGTCGTGCTCCCTTTGCCCAATTACAAGA 426

QY 227 CCTTCGAATCTAACTCAAGAAGCATCAATGACTTTTGACACAAGGTATATAACTA 280
Db 425 CTTTCGATTTCTAACTCAAGAGGATCAATGACTTCGACACGAGGTATATAGCTA 372

RESULT 3
L38243
LOCUS          446 bp mRNA linear EST 03-JUL-1995
DEFINITION    BNAF0581E Mustard flower buds Brassica rapa cDNA, mRNA sequence.
ACCESSION     L38243
VERSION       L38243.1 GI:887283
KEYWORDS      EST.
SOURCE        Brassica rapa (Brassica campestris)
ORGANISM      Brassica rapa
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
               rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE     1 (bases 1 to 446)
AUTHORS       Lim,C.O., Kim,H.Y., Kim,M.G., Lee,S.I., Chung,W.S., Park,S.H.,
               Hwang,I., and Cho,M.J.
TITLE         Expressed sequence tags of Chinese cabbage flower bud cDNA
JOURNAL       Plant Physiol. 111 (2), 577-588 (1996)
PUBMED        8787028
COMMENT       Contact: Lim,C.O., Kim,M.G., Hwang,I. and Cho,M.J.
               Plant Molecular Biology and Biotechnology Research Center
               Email: pmbr@cnongae.gnu.ac.kr.
               Location/Qualifiers
                   1..446
                       /organism="Brassica rapa"
                       /mol_type="mRNA"
                       /strain="pekinensis"
                       /db_xref="taxon:3711"
                       /clone_lib="Mustard flower buds"
                       /note="Devel_stage = flower bud "

ORIGIN
Query Match          36.0%; Score 137.2; DB 8; Length 446;
Best Local Similarity 71.4%; Pred. No. 6.2e-26;
Matches 197; Conservative 0; Mismatches 73; Indels 6; Gaps 1;

QY 12 CTCAGAAGTTCAATCTCACCATGGACGTTTAGTCAAAACAAGATTTTCGAGAGGCGCTT 71
Db 81 CTCAGTTCTATGAGCTCCTCTTGGACGTTCTAAGCAAAATAGATATTCGAAAGGCGCTT 140

QY 72 GGCAGTTTACGACAGGACACACCCGACCGATGGCACATTCGTCGAGGATCTCATCAACAT 131
Db 141 AGCCGTTTACGACAAAGACATCCAGACGTTTGGCAAAACGTCGGCAAGCCGTCGGGAA 200

QY 132 GAAACTGTAGAAGAGTGAAGCGCCACTATGACATTCGTCGAGGATCTCATCAACAT 191
Db 201 TAAATCGGAGAGGAAGTCAAAGCTCACTACGATATTCGTTGAAGATCTCATGAACAT 260

QY 192 CGAGACTGGTCGTGTCCTTTGCCCAATTACAAGACCTTCGA-----ATCTAACTCAAG 245
Db 261 CGAGCAAGACTTAGTCCCTTTGCCTAAATACAAGACCGTCGATGTTGGAAATAAATCAAG 320

QY 246 AGCATCATGACTTTTGACACAGGTATATAACTAA 281
Db 321 AGGCATCAACGGTTACGGTTTGAGGTTTAATGAAAA 356

RESULT 4
CO052301
LOCUS          456 bp mRNA linear EST 14-JUN-2004
DEFINITION    Mdfw2057o18.y1 Mdfw Malus x domestica cDNA clone Mdfw2057o18.5'
               similar to TR:Q9S1J5 Q9S1J5 AT2G21650 PROTEIN.1, mRNA sequence.
ACCESSION     CO052301
VERSION       CO052301.1 GI:48693766

```

KEYWORDS

SOURCE Malus x domestica

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Rosales; Rosaceae; Maloideae; Malus.

REFERENCE

AUTHORS

Korban, S., Vodkin, L., Liu, L., Gasic, K., Gonzales, O., Hernandez, A., Aldwinckle, H., Malnoy, M., Carroll, N., Goldsborough, P., Orvis, K., Clifton, S., Pape, D., Marra, M., Hillier, L., Martin, J., Wylie, T., Dante, M., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Ronko, I., Tsagareishvili, R., Kennedy, S., Waterston, R., and Wilson, R.
Apple Functional Genomics Grant - NSF 0321702
Unpublished (2004)

TITLE JOURNAL COMMENT

Contact: Schuyler S. Korban
Apple Functional Genomics grant - NSF 0321702
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

Library materials provided by: Schuyler S. Korban Library
constructed by: A. Hernandez / K. Gasic Library sequenced by:
Washington University Genome Sequencing Center

WashU EST name: aah28h09.y1

Seq primer: -40UP from Gibco

High quality sequence stop: 456.

Location/Qualifiers

FEATURES

source

1..456
/organism="Malus x domestica"

/mol_type="mRNA"

/cultivar="GoldRush"

/db_xref="taxon:3750"

/clone="WdFw2057018"

/tissue_type="Flower"

/lab_host="DH10B ampicillin resistant"

/clone_lib="WdFw"

/note="Vector: pBluescript II SK (+); Site 1: NotI; Site 2: EcoRI; Total RNA was extracted separately from each floral stage (bud, balloon, open and after pollination), using the 'pine tree' method. Poly(A)+mRNA was isolated twice from total RNA from each stage using the Oligotex Direct mRNA kit (Qiagen). mRNA was reverse transcribed into double stranded cDNA using a modified oligo18(dT) primer with an identifying tag sequence (see table below). cDNA's from different stages were pooled in equal amounts before adaptor ligation. Tag identification when sequencing from 5' end: Stage 1 (bud) insert 18(A)TCGGA; Stage 2 (balloon) insert 18(A)TCGGA; Stage 3 (open) insert 18(A)TCGGT; Stage 4 (after pollination) insert 18(A)TCGGT. Tag identification when sequencing from 3' end: Stage 1 (bud) TCGGA18(T) insert; Stage 2 (balloon) TCGGA18(T) insert; Stage 3 (open) ACGCA18(T) insert; Stage 4 (after pollination) ACGCA18(T) insert. Double stranded cDNAs were size selected (more than 450 bp), adapted with EcoRI adaptors at both ends and then digested with NotI. The cDNAs were then directionally cloned into EcoRI-NotI digested pBS II SK(+) phagemid vector(Stratagene). Identification of adaptors and tags in 5'-end sequenced clones: (Vector)...TAAGCTT(End Vector)(Start EcoRI adaptor)GATATCGAATTCATGTGTGGG (End EcoRI adaptor)(Start Insert)...AAAAAAAAAAAAAAAAA(End Insert)(Start Tag)TCGGA(End Tag)(Start NotI site)/vector/GCGCGCCGCGCGG... The total number of white colony forming units (cfu) in the primary library before amplification was 1.1x10⁶ cfu (colony forming units). The background of empty clones was less than 1%. Inserts ranged from 0.5kb to 3 kb, as determined by PCR. Purified plasmid DNA from the primary library was converted to single-stranded circles and used as a template for PCR amplification using the T7 and T3 priming sites flanking the cloned cDNA inserts. The purified PCR products, representing the entire cloned cDNA population,

were used as a driver for normalization. Hybridization between the single-stranded library and the PCR products was carried out for 44 hours at 30C. Unhybridized single-stranded DNA circles were separated from hybridized DNA rendered partially double-stranded and electroporated into DH10B cells to generate the normalized library. The total number of clones with insert was 9x10⁶ cfu. Background of empty clones was less than 1%."

ORIGIN

Query Match	33.8%	Score 128.8	DB 7	Length 456
Best Local Similarity	69.4%	Pred. No. 1.1e-23		
Matches 175	Conservative 0	Mismatches 77	Indels 0	Gaps 0
QY	3	GSGCGTCAAACTCAAGAGTTCAATCTCACCATTGGACGTTTACTCAAAAAGATGTTTCA	62	
Db	56	GTTTTCTTCTCGAGGAATTTCAACTCTCTGTTGGACGCGGAGGAGAACAAAGCTGTTCA	115	
QY	63	GAGGCGCTTGGCAGTTTACGACAGGACACACCCGACCGATGGCAATGTGGCAAAAGC	122	
Db	116	GAAGCGCTGGCGGTGACGACAGGACAGGACGCGCTGGCATATGTGGCAAGGC	175	
QY	123	TGTCGAGGAGAAACTGTAGAAAGTGAAGCGCACATATGATCTTCGTGAGGATCT	182	
Db	176	CGTCGCGGAAATCTGCGGAGAGTAAGAGGACACTACGAGATCTCGTGCATGATCT	235	
QY	183	CATCAACATCGAGACTGCTGCTCCCTTTGGCCCAATACAGACCTTCGATCTCACTC	242	
Db	236	CATGCACATAGAAATCCGCGCATGTCCTTATACCAATATATAAGTCGCGCGGAGCGG	295	
QY	243	AAGAAGCATCAA	254	
Db	296	CAGCGGACGCAA	307	

RESULT 5

CF230621

LOCUS

DEFINITION

ACCESSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

```

/sex="female"
/tissue_type="cambial zone harvested on the bark side"
/dev_stage="3-years-old poplar trees grown in the nursery"
/clone_lib="Poplar cDNA library from cambial zone"
/notes="A composite cDNA library was made with mRNA
isolated from opposite and tension wood tissues
corresponding to the cambial zone collected on the bark
side after debarking the stem. In this respect, in
addition to cambium cDNA, this library also contains very
young phloem and very young xylem cDNA. The sampling was
done on 3 different tilted trees grown in the nursery.
cDNA were cloned in an oriented way into SfiI (A and B)
restriction sites. A one-step conversion of lambda
triplex to the corresponding triplex2 plasmid was done
via site-specific recombination at loxP sites (Clontech ;
SMART cDNA library construction kit). cDNA inserts were
PCR amplified using flanking primers and then sequenced on
a ABI3100 Genetic Analyser (Applied Biosystem)"

```

ORIGIN

```

Query Match      32.6%; Score 124.2; DB 6; Length 632;
Best Local Similarity 74.6%; Pred. No. 2e-22;
Matches 156; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 19 AGTTCAATCTCACCATTGAGCTTTAGTCAAAAACAAGATGTTCCAGAGGGCCTTGGCAGTT 78
    |||||
Db 211 AATTCTCTCACCTCTTTGGACACCTTAAGCAAAACAAGATCTTCGAAAAGGCCCTGGCTTTA 270
    |||||

QY 79 TAGCACAAGGACACACCCGACCGATGCGCAAAAGCTGTGCGAGGGGAAAACT 138
    |||||
Db 271 TAGCACAAGGACACCCCTGACCCCTGGCATTAATGTTGCCACAGCTGTGGTGGGAATCT 330

QY 139 GTAGAAGAAGTGAAGCGCCACTATGACATTTCTCGTCGAGGATCTCATCAACATCGAGACT 198
    |||||
Db 331 GCAGACGAGTGAAGAGGCACTATGAGATTTCTCATCAAGGATGTCAGGGAATCGAGTCT 390

QY 199 GGTGCGTGCCCTTTGGCCCAATTACAAGAC 227
    |||||
Db 391 GGCAGAGTTCCTTCCTTAATTACAGTTC 419

```

RESULT 6

```

CV277080/c
LOCUS      662 bp      mRNA      linear      EST 22-SEP-2004
DEFINITION WS0142.B21_J13 PTX-D-IL-A-5 Populus trichocarpa x Populus deltoides
            cDNA clone WS0142_J13 3', mRNA sequence.
CV277080
VERSION    CV277080.1 GI:52530055
KEYWORDS   EST.
SOURCE     Populus trichocarpa x Populus deltoides
ORGANISM   Populus trichocarpa x Populus deltoides

```

```

Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.
1 (bases 1 to 662)
Ralph,S., Cooper,D., Kolosova,N., Oddy,C., Butterfield,Y.,
Kirkpatrick,R., Liu,J., Palmquist,D., Stott,J., Barber,S., Yang,G.,
Babakaiff,R., Brown-John,M., Chand,S., Featherstone,R., Maason,A.,
Mayo,M., Moran,J., Olson,T., Wong,D., Ritland,C.E., Siddiqui,A.,
Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C., Ritland,K. and
Bohlmann,J.

```

REFERENCE

```

AUTHORS    The poplar transcriptome: Analysis of expressed sequence tags from
            multiple cDNA libraries
JOURNAL     Unpublished (2004)
COMMENT     Contact: Joerg Bohlmann
            Genome BC forest genomics program
            University of British Columbia
            Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
            Vancouver, British Columbia, Canada, V6T 1Z3
            Tel: 1-604-822-0282
            Fax: 1-604-822-2114
            Email: bohlmann@msl.ubc.ca
            Plate: WS0142 row: J column: 13

```

```

High quality sequence stop: 662
POLYA=yes.

```

FEATURES

source

```

Location/Qualifiers
1..662
/organism="Populus trichocarpa x Populus deltoides"
/mol_type="mRNA"
/cultivar="H11-11"
/db_xref="taxon:3695"
/clone="WS0142_J13"
/sex="Male"
/lab_host="E. coli DH10B T1 phage resistant cells"
/clone_lib="PTXD-IL-A-5"
/notes="Vector: pBluescript II SK (+) XR; Site 1: EcoRI (5'
end of cDNA); Site 2: XhoI (3' end of cDNA); Sapling trees
two metres in height and grown under greenhouse conditions
were exposed to continuous feeding by Malacosoma disstria
Hubner (forest tent caterpillar) mid-instar larvae caged
on the sapling using mesh bags. Mature leaves from within
the caged region were collected 2 hours, 12 hours, 24
hours and 48 hours after the onset of treatment. mRNA was
isolated from each tissue source independently and equal
quantities of mRNA from each tissue were then pooled. cDNA
was prepared from 5 micrograms of mRNA and directionally
ligated into the pBluescript II SK (+) XR vector using the
pBluescript II XR cDNA Library Construction kit according
to manufacturer's instructions with modifications
(Stratagene). Plasmid DNA was then transformed by
electroporation into DH10B cells (Invitrogen) for
propagation."

```

ORIGIN

```

Query Match      32.2%; Score 122.6; DB 7; Length 662;
Best Local Similarity 74.2%; Pred. No. 5.5e-22;
Matches 155; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 19 AGTTCAATCTCACCATTGAGCTTTAGTCAAAAACAAGATGTTCCAGAGGGCCTTGGCAGTT 78
    |||||
Db 481 AATTCTCTCACCTCTTTGGACACCTTAAGCAAAACAAGATTCGAAAAGGCCCTGGCTTTA 422

QY 79 TAGCACAAGGACACACCCGACCGATGCGCAAAAGCTGTGCGAGGAGGAAAACT 138
    |||||
Db 421 TATGACAGGACACCCCTGACCGCTGGCATGATGTTGCCAAAAGCTGTGGTGGGAATCT 362

QY 139 GTAGAAGAAGTGAAGCGCCACTATGACATTTCTCGTCGAGGATCTCATCAACATCGAGACT 198
    |||||
Db 361 GCAGAGGAAGTGAAGAGGCACTATGAGATTTCTCATCAAGATGTCAGGGAATCGAGTCT 302

QY 199 GGTGCGTGCCCTTTGGCCCAATTACAAGAC 227
    |||||
Db 301 GGCAGAGTTCCTTCCTTAATTACAGTTC 273

```

RESULT 7

```

CV277080/c
LOCUS      728 bp      mRNA      linear      EST 17-OCT-2002
DEFINITION P040E03 Populus petioles cDNA library Populus tremula cDNA 5 prime,
            mRNA sequence.
CV277080
VERSION    CV277080.1 GI:24101759
KEYWORDS   EST.
SOURCE     Populus tremula
ORGANISM   Populus tremula

```

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.
1 (bases 1 to 728)
Unneberg,P., Bhalerao,R.R., Jansson,S. and Sterky,F.
The poplar tree transcriptome: Analysis of expressed sequence tags
from multiple libraries
Unpublished (2002)
JOURNAL     Contact: BHALERAO RUPALI R.
            Umea Plant Science Center
            Department of Plant Physiology

```


University of Umea, 901 87 Umea, Sweden
Tel: +46 90 786 5279
Fax: +46 90 786 6676
Email: rupali.bhalerao@plantphys.umu.se.

FEATURES

source
Location/Qualifiers
1..728
/organism="Populus tremula"
/mol_type="mRNA"
/db_xref="taxon:113636"
/tissue_type="petioles"
/clone_lib="Populus petioles cdna library"

ORIGIN

Query Match 32.0%; Score 121.8; DB 5; Length 728;
Best Local Similarity 68.6%; Pred. No. 9.2e-22;
Matches 168; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 13 TCAAGAGTTCAATCTCACCATTGGAGCTTTAGTCAAAACAAGATGTCGAGAGGCGCTTG 72
DB 257 TCAAAATGGCTCGGCTCTCTCTGGACAGCCAAACAAAACAAGCTATTTCGAGAAGGCCCTG 316

QY 73 GCAGTTTACGACAAGACACACCCGACGATGCGCAATGTGCAAAAGCTCTCGAGGG 132
DB 317 GCTGTATACGACAAGACACCCGACGCTGGCAAAATGTGCCAAGCCGCTGGTGC 376

QY 133 AAACTGTAGAGAAGTGAAGCGCCACTATGACATTTCTGTCGAGGATCTCATCAATC 192
DB 377 AAGTCTCTCGAAGAAGTTAAGAGGCACTATGATCTGTCGAGAGTCTGCTGACATA 436

QY 193 GAGACTGTGCTGCTCTCTTGGCCAAATCAAGACCTTCGAATCTTAATCAAGAAGCATC 252
DB 437 GAATCGGCGAGCGCTCTGCGAATTACAGCCCTCTGGCAGCAATGGTAGAGGACTT 496

QY 253 AATGA 257
DB 497 GGTGA 501

RESULT 8
CD669972 516 bp mRNA linear EST 23-JUN-2003
LOCUS eepic.pk005.g21 eepic Eucalyptus grandis cdna clone eepic.pk005.g21
DEFINITION 5' end, mRNA sequence.

ACCESSION CD669972
VERSION CD669972.1 GI:32164642

KEYWORDS EST
SOURCE Eucalyptus grandis

ORGANISM

Eucalyptus grandis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; Myrtales; Myrtaceae; Eucalyptus.

REFERENCE

AUTHORS Tingey,S.V., Dolan,M., Sedoreff,R.R., Hanafey,M.K., Kirst,M.,
Myburg,A.A. and Egertsdotter,U.

TITLE

Eucalyptus EST Project

JOURNAL

Comment: Scott V. Tingey

COMMENT

E. I. DuPont de Nemours and Company
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
Tel: 302-631-2602
Fax: 302-631-2607
Email: Scott.V.Tingey@usa.dupont.com

Seq primer: M13.

FEATURES

source
Location/Qualifiers
1..457
/organism="Eucalyptus grandis"
/mol_type="mRNA"
/db_xref="taxon:71139"
/clone="eepic.pk005.g21"
/tissue_type="leaf"
/lab_host="DH10B"
/clone_lib="eepic"

ORIGIN

Query Match 31.8%; Score 121.2; DB 6; Length 457;
Best Local Similarity 70.4%; Pred. No. 1.2e-21;
Matches 162; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 28 TCACCATGCGCTTTAGTCAAAACAAGATGTCGAGAGGCGCTTGGCAGTTTACGACAAG 87
DB 52 TCATCTGGACGCCGAGCAAGATGTTTCGAGAAGGCAATTGGCTCAATATGACAAG 111

QY 88 GACACACCCGACCGATGGCACAATGTGCAAAAGCTGTCGAGAGGAAAACCTGTAGAAGAA 147
DB 112 GACATCCCGACAGGTGGCAGAAGATTGCCAAGGCCGTGGTGGGAAATCTGCAGATGAA 171

QY 148 GTGAAGCCGCACTATGACATTTCTGTCGAGGATCTCATCAACATCGAGACTGGTGGTGC 207
DB 172 GTGAAAAGACACTATGAAAATTTTATCGAGGACGTCGAAGCACATCGAGTCTGCAGAGTT 231

QY 208 CCTTTGCCCAATTAACAAGACCTTCGAATCTAACTCAAGAAGCATCAATGA 257
DB 232 CCTTTTCTTAATTACAGTTCGAGCAACAATAGTGGTGAAGGTTCCGGTGA 281

RESULT 9

CV278823/c 516 bp mRNA linear EST 22-SEP-2004
LOCUS WS0147.B21.111 PTxD-IL-A-5 Populus trichocarpa x Populus deltoides
DEFINITION cDNA clone WS0147_111 3', mRNA sequence.

ACCESSION

CV278823

VERSION

CV278823.1 GI:52531798

KEYWORDS

EST.

SOURCE

Populus trichocarpa x Populus deltoides
Populus trichocarpa x Populus deltoides
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.

REFERENCE

AUTHORS Ralph,S., Cooper,D., Kolosova,N., Oddy,C., Butterfield,Y.,
Kirkpatrick,R., Liu,J., Palmquist,D., Scott,J., Barber,S., Yang,G.,
Babakaiff,R., Brown-John,M., Chand,S., Featherstone,R., Masson,A.,
Mayo,M., Moran,J., Olson,T., Wong,D., Ritland,C.E., Siddiqui,A.,
Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C., Ritland,K. and
Bohlmann,J.

The poplar transcriptome: Analysis of expressed sequence tags from
multiple cDNA libraries
Unpublished (2004)

JOURNAL

COMMENT

Contact: Joerg Bohlmann
Genome BC forest genomics program
University of British Columbia
Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
Vancouver, British Columbia, Canada, V6T 1Z3
Tel: 1-604-822-0282
Fax: 1-604-822-2114

Email: bohlmann@msl.ubc.ca

Plate: WS0147 row: 1 column: 11

High quality sequence stop: 516

POLYA=yes.

FEATURES

Location/Qualifiers

1..516

source

/organism="Populus trichocarpa x Populus deltoides"
/mol_type="mRNA"
/cultivar="H11-11"

/db_xref="taxon:3695"

/clone="WS0147_111"

/sex="Male"

/lab_host="E. coli DH10B T1 phage resistant cells"

/clone_lib="PTxD-IL-A-5"

/note="Vector: pBluescript II SK (+) XR; Site 1: EcoRI; Site 2: XhoI (3' end of cDNA); Sapling trees two metres in height and grown under greenhouse conditions were exposed to continuous feeding by Malacosoma disstria

Db 389 ACATAGAATCGGCCAAGCCCTCTCGCGAATTACAAGCCCTTCGAGCAATGGTAGAG 330
 QY 248 GCATCAATGA 257
 Db 329 GACTTGTGA 320

RESULT 13

CB920052 626 bp mRNA linear EST 25-APR-2003
 LOCUS VVD058F11 350737 An expressed sequence tag database for abiotic
 DEFINITION stressed berries of Vitis vinifera var. Chardonnay Vitis vinifera
 CDNA clone VVD058F11 5, mRNA sequence.

ACCESSION CB920052
 VERSION
 KEYWORDS EST.

SOURCE Vitis vinifera
 ORGANISM Vitis vinifera
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; Vitaceae; Vitis.
 1 (bases 1 to 626)

REFERENCE Cushman,J.C.
 AUTHORS An expressed sequence tag database for abiotic stressed berries of
 TITLE Vitis vinifera var. Chardonnay

JOURNAL Vitis vinifera var. Chardonnay
 COMMENT Unpublished (2002)
 Contact: Cushman JC
 Department of Biochemistry
 University of Nevada
 MS200, Reno, NV 89557-0014, USA
 Tel: 775-784-1918
 Fax: 775-784-1650
 Email: jcushman@unr.edu

PCR Primers
 FORWARD: T3 20mer
 BACKWARD: T7 21mer (backward)
 Plate: 058 row: F column: 11
 Seq primer: T3 20mer
 High quality sequence stop: 626.

FEATURES

source
 1..626
 /organism="Vitis vinifera"
 /mol_type="mRNA"
 /db_xref="taxon:29760"
 /clone="VVD058F11"
 /tissue_type="berries"
 /dev_stage="mixed; 8, 9, 11, 13, 15, 16 weeks daf"
 /clone_lib="An expressed sequence tag database for abiotic
 stressed berries of Vitis vinifera var. Chardonnay"
 /note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site_1:
 EcoRI; Site_2: XhoI"

ORIGIN

Query Match 31.4%; Score 119.8; DB 6; Length 626;
 Best Local Similarity 70.5%; Pred. No. 3.1e-21;
 Matches 160; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 12 CTCAGAAGTTCAATCTCACCATGGACGTTTAGTCAAAACAAGATGTCGAGAGGGCCTT 71
 |||||
 Db 168 CTCGCCAGCTCTGGCTCATCTCTGGACTGCCAAGCAAAACAAGCCTTTGAAGAGGGCCT 227

QY 72 GGCAGTTTACGACAAGGACACACCCGATGGCACAATGTGGCAAAAGCTGTTCGAGG 131
 |||||
 Db 228 TGCAGTGTATGACAAGGACACCCCTGACCGTGTGTACAATGTTGCCAGGGCTGTGTGG 287

QY 132 GAAACCTGTAGAAGTGAAGCCCACTATGACATTCCTGTCAGAGATCTCATCAAT 191
 |||||
 Db 288 GAAACCCGTGGAGGAAGTGAAGGGCACTATGAGATCCTTTGTGAGGATATCAAGAGCAT 347

QY 192 CGAGACTGGTCTGCTCCCTTTGCCCAATTACAAGACCTTCGAATCTA 238
 |||||
 Db 348 CGACTCAGACAAAGTGGCCCTTCCCAATTACAAGACTACCGGAGCTA 394

RESULT 14

CB919185
 LOCUS
 DEFINITION

CB919185 662 bp mRNA linear EST 25-APR-2003
 VVD042D03 349003 An expressed sequence tag database for abiotic
 stressed berries of Vitis vinifera var. Chardonnay Vitis vinifera
 CDNA clone VVD042D03 5, mRNA sequence.

ACCESSION CB919185
 VERSION
 KEYWORDS EST.

SOURCE Vitis vinifera
 ORGANISM Vitis vinifera

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; Vitaceae; Vitis.
 1 (bases 1 to 662)

REFERENCE Cushman,J.C.
 AUTHORS An expressed sequence tag database for abiotic stressed berries of

TITLE Vitis vinifera var. Chardonnay

JOURNAL Unpublished (2002)

COMMENT Contact: Cushman JC
 Department of Biochemistry
 University of Nevada
 MS200, Reno, NV 89557-0014, USA
 Tel: 775-784-1918
 Fax: 775-784-1650
 Email: jcushman@unr.edu

PCR Primers
 FORWARD: T3 20mer
 BACKWARD: T7 21mer (backward)
 Plate: 042 row: D column: 03
 Seq primer: T3 20mer
 High quality sequence stop: 662.

Location/Qualifiers
 1..662

/organism="Vitis vinifera"
 /mol_type="mRNA"
 /db_xref="taxon:29760"

/clone="VVD042D03"
 /tissue_type="berries"

/dev_stage="mixed; 8, 9, 11, 13, 15, 16 weeks daf"
 /clone_lib="An expressed sequence tag database for abiotic
 stressed berries of Vitis vinifera var. Chardonnay"

/note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site_1:
 EcoRI; Site_2: XhoI"

Query Match 31.4%; Score 119.8; DB 6; Length 662;
 Best Local Similarity 70.5%; Pred. No. 3.1e-21;
 Matches 160; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 12 CTCAGAAGTTCAATCTCACCATGGACGTTTAGTCAAAACAAGATGTCGAGAGGGCCTT 71
 |||||
 Db 168 CTCGCCAGCTCTGGCTCATCTCTGGACTGCCAAGCAAAACAAGCCTTTGAAGAGGGCCT 227

QY 72 GGCAGTTTACGACAAGGACACACCCGATGGCACAATGTGGCAAAAGCTTCGAGG 131
 |||||
 Db 228 TGCAGTGTATGACAAGGACACCCCTGACCGTGTGTACAATGTTGCCAGGGCTGTGTGG 287

QY 132 GAAACCTGTAGAAGTGAAGCCCACTATGACATTCCTGTCAGAGATCTCATCAAT 191
 |||||
 Db 288 GAAACCCGTGGAGGAAGTGAAGGGCACTATGAGATCCTTTGTGAGGATATCAAGAGCAT 347

QY 192 CGAGACTGGTCTGCTCCCTTTGCCCAATTACAAGACCTTCGAATCTA 238
 |||||
 Db 348 CGACTCAGACAAAGTGGCCCTTCCCAATTACAAGACTACCGGAGCTA 394

RESULT 15
 BU668323
 LOCUS

DEFINITION
 MC01026A10 MC01 Sesamum indicum cDNA, mRNA sequence.

ACCESSION
 BU668323

```

VERSION      BU668323.1  GI:40446455
KEYWORDS     EST.
SOURCE       Sesamum indicum (sesame)
ORGANISM     Sesamum indicum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Lamiales; Pedaliaceae; Sesamum.
REFERENCE    1 (bases 1 to 474)
AUTHORS      Suh,M.C.
TITLE        Comparative analysis of Expressed Sequence Tags between Sesamum
JOURNAL      indicum and Arabidopsis thaliana developing seeds
COMMENT      Unpublished (2003)
              Contact: Suh, Mi Chung
              Graduate School of Biotechnology, Korea University
              1, 5-Ka, Anam-dong, Sungbuk-ku, Seoul 136-701, Rep. of Korea
              Tel: 82 2 3290 3169
              Fax: 82 2 927 9028
              Email: michung@korea.ac.kr
              Plate: 026 row: A column: 10.
              Location/Qualifiers
                1..474
                  /organism="Sesamum indicum"
                  /mol_type="mRNA"
                  /db_xref="taxon:4182"
                  /tissue_type="developing seed"
                  /cell_line="Yangbaeck"
                  /dev_stage="5 to 25 days after flowering"
                  /clone_lib="MC01"
                  /note="Vector: pBluescript SK; Site_1: EcoRI; Site_2:
XhoI"

FEATURES
    source
        1..474
            /organism="Sesamum indicum"
            /mol_type="mRNA"
            /db_xref="taxon:4182"
            /tissue_type="developing seed"
            /cell_line="Yangbaeck"
            /dev_stage="5 to 25 days after flowering"
            /clone_lib="MC01"
            /note="Vector: pBluescript SK; Site_1: EcoRI; Site_2:
XhoI"

ORIGIN
Query Match      31.0%; Score 118.2; DB 5; Length 474;
Best Local Similarity 67.9%; Pred. No. 8e-21;
Matches 165; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY      6  GTCAAACTCAAGAAGTTCAATCTCACCATGGACGTTTAGTCAAAAACAAGATGTTGAGAG 65
Db      179  GACATCCGCACGTGGCTCAGCCTCAGCGTGGAGCTCTAAGCAAAACAAGCAGTTCGAGGA 238

QY      66  GGCCTTGGCAGTTTACGACAAGGACACACCCGACCGATGGCACATGTGGCAAAAGCTGT 125
Db      239  GGCCTTGGCAATGTACGACAAGGACACGCGCTGACCGTTGGCATAATATAGCCCGAGCCGT 298

QY      126  CGGAGGGAAACTGTAGAGAAGTGAAGCGCCACTATGACATTCTCGTCGAGGATCTCAT 185
Db      299  GGGTGGGAAGTCAGCAGAGGAAGTGAGAAGGCATTATGAGATTCTGGTTAAGGACATTAT 358

QY      186  CAACATCGAGACTGGTGTGTCCTTTGCCCCAATTACAGACCTTCGAATCTTAAGTCAAG 245
Db      359  GCAGATCGAGACGGATCAGGTGCCAATACCTTAACCTACAGAGCCATGGGGAGCAACAGCAG 418

QY      246  AAG 248
Db      419  AGG 421

```

Search completed: March 10, 2006, 19:30:36
Job time : 2651 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 10, 2006, 18:28:27 ; Search time 400 Seconds
(without alignments)
6348.120 Million cell updates/sec

Title: US-10-697-787-1

Perfect score: 381

Sequence: 1 atggcgctcaactcaagaag.....gttatattcttggtttaa 381

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*
- 14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	381	100.0	381	10 ADF43565	Adf43565 Thale cre
2	273.4	71.8	294	14 AEA27134	Aea27134 Stress to
3	149.4	39.2	279	12 ADO62042	Ado62042 Transcrip
4	138.6	36.4	478	3 AAC34192	Aac34192 Arabidops
5	122.8	32.2	482	14 ADW16587	Adw16587 Eucalyptu
6	120.6	31.7	760	13 ADR59489	Adr59489 Cotton cd
7	119	31.2	543	13 ACN61386	Acn61386 Cotton gy
8	110.8	29.1	779	13 ADR59490	Adr59490 Cotton cd
9	110.8	28.9	466	12 ADJ43786	Adj43786 Plant cdn
10	108.6	28.5	513	11 ACL30919	Acl30919 Rice ablo
11	108.4	28.5	508	3 AAC35256	Aac35256 Arabidops
12	108.4	28.5	577	13 ACN59055	Acn59055 Cotton gy
13	107.4	28.2	505	13 ACN61224	Acn61224 Cotton gy
14	107.4	28.2	532	13 ACN59701	Acn59701 Cotton gy
15	107.4	28.2	532	13 ACN56695	Acn56695 Cotton gy
16	107.4	28.2	538	13 ACN56695	Acn56695 Cotton gy
17	107.4	28.2	542	13 ACN56782	Acn56782 Cotton gy
18	107.4	28.2	604	13 ADR64219	Adr64219 Cotton cd
19	104.8	27.5	306	3 AAC42715	Aac42715 Arabidops

c	20	104.8	27.5	556	10 ABX57121	Abx57121 Arabidops
	21	104.8	27.5	578	10 ADD30327	Add30327 Plant vie
	22	104.8	27.5	578	12 ADI44256	Adi44256 Plant tra
	23	104.8	27.5	578	14 AEA27034	Aea27034 Stress to
	24	103.8	27.2	636	13 ADR65153	Adr65153 Cotton cd
	25	103.4	27.1	584	13 ACN54517	Acn54517 Cotton an
	26	99	26.0	504	3 AAC49819	Aac49819 Arabidops
	27	98.6	25.9	507	13 ADX14446	Adx14446 Plant ful
	28	92.4	24.3	513	10 ADB31830	Adb31830 DNA encod
	29	92.4	24.3	513	12 ADO02248	Ado02248 Thalecres
	30	92.4	24.3	513	14 AEA27102	Aea27102 Stress to
	31	92	24.1	294	3 AAC43000	Aac43000 Arabidops
	32	92	24.1	294	14 AEA27038	Aea27038 Stress to
	33	92	24.1	303	3 AAC46353	Aac46353 Arabidops
	34	91	23.9	2192	3 AAA78533	Aaa78533 Plant SDF
	35	87	22.8	980	13 ADX51987	Adx51987 Plant ful
	36	84.6	22.2	776	12 ADJ39433	Adj39433 Plant cdn
c	37	83.6	21.9	536	13 ACN58748	Acn58748 Cotton gy
	38	81.2	21.3	563	13 ACN58701	Acn58701 Cotton gy
	39	79.6	20.9	867	6 ABK65275	Abk65275 Arabidops
	40	79.6	20.9	867	10 ADD30621	Add30621 Plant yie
	41	79.6	20.9	867	10 ADE31552	Adel31552 Plant yie
	42	79.6	20.9	867	12 ADI44216	Adi44216 Plant tra
	43	79.4	20.8	1431	14 ADW16598	Adw16598 Eucalyptu
	44	77.6	20.4	1230	13 ADX13094	Adx13094 Plant ful
	45	75	19.7	633	13 ADX12277	Adx12277 Plant ful

ALIGNMENTS

RESULT 1
ADF43565
ID ADF43565 standard; DNA; 381 BP.
XX
AC ADF43565;
XX
DT 12-FEB-2004 (first entry)
XX
DE Thale cress peronospora parasitica resistant protein coding sequence.
XX
KW transgenic plant; peronospora parasitica resistant; PPR2;
KW pathogen resistant phenotype; modified pathogen resistance; thale cress;
KW gene; ds.
XX
OS Arabidopsis thaliana.
XX
FH Key Location/Qualifiers
CDS 1..381
FT /*tag= a
FT /product= "Thale cress peronospora parasitica resistant
FT protein"
XX
PN WO2003091412-A2.

XX
PD 06-NOV-2003.
XX
PF 24-APR-2003; 2003WO-US012981.
XX
PR 24-APR-2002; 2002US-0375333P.
XX
(AGRI-) AGRINOMICS LLC.
XX
PI Federspiel N, Lammers A, Liu XL, Bates SR, Westerlund C;
PI Fitch JR;
XX
WPI; 2003-865582/80.
P-PSDB; ADF43566.

XX
PT New transgenic plants with increased resistance to pathogens due to
PT altered expression of Peronospora Parasitica Resistant gene (PPR2),
PT useful for generating plants with a pathogen resistance phenotype.
XX

```
PS Example 4; SEQ ID NO 1; 36pp; English.
XX
CC The invention relates to a transgenic plant that possesses a plant
CC transformation vector comprising a nucleotide sequence that encodes a
CC peronospora parasitica resistant (PPR2) protein, or a PPR2 orthologue.
CC The transgenic plant is useful in generating plants with a pathogen
CC resistance phenotype. The PPR2 nucleic acids and proteins of the
CC invention are useful in the generation of genetically modified plants
CC having a modified pathogen resistance phenotype. The present DNA sequence
CC encodes an Arabidopsis thaliana PPR2 protein of the invention.
XX
SQ Sequence 381 BP; 122 A; 76 C; 76 G; 107 T; 0 U; 0 Other;

Query Match 100.0%; Score 381; DB 10; Length 381;
Best Local Similarity 100.0%; Pred. No. 1.2e-98;
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGTCAAACTCAAGAAGTTCAATCTCACCATGGACGTTTAGTCAAAAACAAGATGTTTC 60
Db 1 ATGGCGTCAAACTCAAGAAGTTCAATCTCACCATGGACGTTTAGTCAAAAACAAGATGTTTC 60
QY 61 GAGAGGGCCCTTGGCAGTTTACGACAAGGACACACCCGACCGATGGCACAATGTGGCAAAA 120
Db 61 GAGAGGGCCCTTGGCAGTTTACGACAAGGACACACCCGACCGATGGCACAATGTGGCAAAA 120
QY 121 GCTGTGCGAGGGAAAAGTGTAGAAGAGTGAAGCGGCCACTATGACATTCCTGTCGAGGAT 180
Db 121 GCTGTGCGAGGGAAAAGTGTAGAAGAGTGAAGCGGCCACTATGACATTCCTGTCGAGGAT 180
QY 181 CTCATCAACATCGAGACTGGTGTGTCCTTTGCCCAATTACAGACCTTCGAAATCTTAAC 240
Db 181 CTCATCAACATCGAGACTGGTGTGTCCTTTGCCCAATTACAGACCTTCGAAATCTTAAC 240
QY 241 TCAAGAAGCATCAATGACTTTTGACACAAGGTATATACTAAATATCTATATATGATGCTC 300
Db 241 TCAAGAAGCATCAATGACTTTTGACACAAGGTATATACTAAATATCTATATATGATGCTC 300
QY 301 TCGATATATTTTGATTAATCATCTTAGTGATTTTGAGAAATCTCTCAAAAAGTTCTTGTA 360
Db 301 TCGATATATTTTGATTAATCATCTTAGTGATTTTGAGAAATCTCTCAAAAAGTTCTTGTA 360
QY 361 AGTTATATTTCTTTGGTTTAA 381
Db 361 AGTTATATTTCTTTGGTTTAA 381

RESULT 2
AEA27134
ID AEA27134 standard; DNA; 294 BP.
XX
AC AEA27134;
XX
DT 28-JUL-2005 (first entry)
XX
DE Stress tolerant plant-related transcription factor gene SeqID975.
XX
KW transcription factor; transgenic plant; agriculture; drought resistance;
KW stress tolerance; gene; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO2005047516-A2.
XX
PD 26-MAY-2005.
XX
PF 12-NOV-2004; 2004WO-US037584.
XX
PR 13-NOV-2003; 2003US-00714887.
PR 05-DEC-2003; 2003US-0527658P.
PR 05-FEB-2004; 2004US-0542928P.
XX
PA (MEND-) MENDEL BIOTECHNOLOGY INC.
XX

PI Heard JE, Riechmann JL, Creelman RA, Ratcliffe OJ, Canales RD;
PI Repetti P, Kunimoto RW, Gutterson NI, Reuber TL, Pineda O;
PI Sherman BK, Morrison TA, Reddie JS, Jiang C, Century KS, Adam L;
PI Zhang JZ, Hempel FD, Libby JM;
XX
DR WPI: 2005-372386/38.
DR P-PSDB; AEA27135.
XX
PT New transgenic plants for producing commercially or agriculturally useful
PT plants having improved tolerance to drought, shade and low nitrogen
PT conditions.
XX
PS Disclosure; SEQ ID NO 975; 407pp; English.
XX
CC This invention relates to a novel plant transcription factor
CC polypeptides, the DNA sequences which encode them and their use in
CC creating transgenic plants. The transgenic plant and methods are useful
CC for producing commercially or agriculturally useful plants having
CC improved tolerance to drought, shade and low nitrogen conditions when
CC compared to wild-type reference plants. The present sequence is that of a
CC plant transcription factor gene which was used during the development of
CC the transgenic plants of the invention.
XX
SQ Sequence 294 BP; 99 A; 65 C; 67 G; 63 T; 0 U; 0 Other;

Query Match 71.8%; Score 273.4; DB 14; Length 294;
Best Local Similarity 97.9%; Pred. No. 6.2e-68;
Matches 277; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ATGGCGTCAAACTCAAGAAGTTCAATCTCACCATGGACGTTTAGTCAAAAACAAGATGTTTC 60
Db 1 ATGGCGTCAAACTCAAGAAGTTCAATCTCACCATGGACGTTTAGTCAAAAACAAGATGTTTC 60
QY 61 GAGAGGGCCCTTGGCAGTTTACGACAAGGACACACCCGACCGATGGCACAATGTGGCAAAA 120
Db 61 GAGAGGGCCCTTGGCAGTTTACGACAAGGACACACCCGACCGATGGCACAATGTGGCAAAA 120
QY 121 GCTGTGCGAGGGAAAAGTGTAGAAGAGTGAAGCGGCCACTATGACATTCCTGTCGAGGAT 180
Db 121 GCTGTGCGAGGGAAAAGTGTAGAAGAGTGAAGCGGCCACTATGACATTCCTGTCGAGGAT 180
QY 181 CTCATCAACATCGAGACTGGTGTGTCCTTTGCCCAATTACAGACCTTCGAAATCTTAAC 240
Db 181 CTCATCAACATCGAGACTGGTGTGTCCTTTGCCCAATTACAGACCTTCGAAATCTTAAC 240
QY 241 TCAAGAAGCATCAATGACTTTTGACACAAGGTATATACTAAAT 283
Db 241 TCAAGAAGCATCAATGACTTTTGACACAAGGAAAATGAAGAATT 283

RESULT 3
ADO62042
ID ADO62042 standard; DNA; 279 BP.
XX
AC ADO62042;
XX
DT 15-JUL-2004 (first entry)
XX
DE Transcription factor G2723 coding sequence, SEQ ID 509.
XX
KW Plant; transcription factor; transgenic plant; abiotic stress tolerance;
KW osmotic stress tolerance; cold tolerance; heat tolerance;
KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
KW glyphosate resistance; flowering; fertility; seed development; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO2004031349-A2.
XX
PD 15-APR-2004.
XX
PF 18-SEP-2003; 2003WO-US030292.
XX
```


PR 18-SEP-2002; 2002US-0411837P.
 PR 17-DEC-2002; 2002US-0434166P.
 PR 24-APR-2003; 2003US-0465809P.
 XX (MEND-) MENDEL BIOTECHNOLOGY INC.
 XX Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;
 PI Riechmann JL, Haake V, Dubell AN, Keddle JS, Sherman BK;
 XX WPI: 2004-330163/30.
 DR P-PSDB; AD062043.
 XX
 XX New recombinant polynucleotide encoding transcription factor
 PT polypeptides, useful for producing transgenic plants with advantageous
 PT properties compared to a reference plant.
 XX
 XX Claim 1; SEQ ID NO 509; 510pp; English.
 XX
 CC The present invention relates to novel plant transcription factor
 CC proteins (I) and nucleotide sequences (II) (AD061534-AD063778). The
 CC sequences can be used to produce transgenic plants, which overexpress
 CC (II), where the transgenic plant has an altered trait as compared to a
 CC non-transgenic plant or wild-type plant. The transgenic plant comprises
 CC an altered trait selected from increased tolerance to abiotic stresses,
 CC increased tolerance to osmotic stress, increased tolerance to cold,
 CC increased germination in cold, increased tolerance to heat, increased
 CC germination in heat, increased tolerance to freezing conditions,
 CC increased tolerance to low nitrogen conditions, increased tolerance to
 CC low phosphate conditions, increased tolerance to disease, including
 CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,
 CC increased tolerance to multiple fungal pathogens, increased resistance to
 CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,
 CC increased sensitivity to ACC, altered sugar sensing, increased tolerance
 CC to sugars, altered carbon/nitrogen sensing, early flowering, late
 CC flowering, altered flower structure, loss of flower determinacy, reduced
 CC fertility, altered shoot meristem development, altered branching pattern,
 CC altered stem morphology, altered vascular tissue structure, reduced
 CC apical dominance, altered trichome density, altered trichome development,
 CC altered trichome structure, altered root development, altered shade
 CC avoidance, altered seed development, altered seed ripening, altered seed
 CC germination, slow growth, fast growth, altered cell differentiation,
 CC altered cell proliferation, altered cell expansion, altered phase change,
 CC altered senescence, abnormal embryo development, altered programmed cell
 CC death, lethality when overexpressed, altered necrosis patterns, increased
 CC plant size, increased biomass, large seedlings, dwarfed plants, dark
 CC green leaves, change in leaf shape, increased leaf size and mass, light
 CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
 CC altered seed coloration, altered seed size, altered seed shape, large
 CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil
 CC content, altered seed protein content, altered seedprenyl content,
 CC altered leaf prenol lipid content, increased anthocyanin levels, and
 CC decreased anthocyanin levels. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 279 BP; 86 A; 56 C; 65 G; 72 T; 0 U; 0 Other;

Query Match 39.2%; Score 149.4; DB 12; Length 279;
 Best Local Similarity 74.0%; Pred. No. 1.7e-32;
 Matches 205; Conservative 0; Mismatches 66; Indels 6; Gaps 1;

QY 1 ATGGCGTCAAACTCAAGATTTCATCTCACCATGGAGCTTTAGTCAAAACAGATGTC 60
 DB 1 ATGGCGTCTAGTTCTATGAGCTCGAGCTTCTTGGAGCTCTAAGCAAAACAGATGTC 60
 QY 61 GAGAGGGCTTTGGCAGTTTACGACAAAGGACACACCCGACCGATGGCAATGTGGCAAAA 120
 DB 61 GAGAGGGCTTTAGCCGTTTACGATAAGACACTCCCGACCGTTGGCAAAACGTCGTAA 120
 QY 121 GCTGTGGAGGAAACATGTGAGAGAGTGAAGGCCCACTATGACATTTCTCGTCAGAT 180
 DB 121 GCAGTTGGAAGTAAATCTGCAGAGGAGTAAACGTCACGTACGACATCCTCGTTGAAGAT 180

QY 181 CTCATCAACATCGAGACTGGTGGTCCCTTTCCCTTTCCCAATTACAAGACCTTGA-----A 234
 DB 181 CTCATGAACATCGAACAAGACTTTAGTACCTTTGGCTTAAATACAAAACCGTCGATGTGGA 240
 QY 235 TCTAACTCAAGAAGCATCAATGACTTTGACACAAGGT 271
 DB 241 AGTAAATCTAGAGGATCGATGATTTGATTGAGGT 277
 RESULT 4
 AAC341192
 ID AAC341192 standard; DNA; 478 BP.
 AC AAC341192;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 5783.
 KW Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway; metabolic pathway;
 KW promoter; termination sequence; ss.
 XX Arabidopsis thaliana.
 OS Arabidopsis thaliana.
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 XX 25-FEB-2000; 2000EP-00301439.
 PR 25-FEB-1999; 99US-0121825P.
 PR 05-MAR-1999; 99US-0123180P.
 PR 23-MAR-1999; 99US-0123548P.
 PR 29-MAR-1999; 99US-0125788P.
 PR 01-APR-1999; 99US-0126264P.
 PR 01-APR-1999; 99US-0126785P.
 PR 06-APR-1999; 99US-0127462P.
 PR 08-APR-1999; 99US-0128234P.
 PR 16-APR-1999; 99US-0128714P.
 PR 19-APR-1999; 99US-0129845P.
 PR 21-APR-1999; 99US-0130077P.
 PR 23-APR-1999; 99US-0130449P.
 PR 28-APR-1999; 99US-0130510P.
 PR 30-APR-1999; 99US-0130891P.
 PR 30-APR-1999; 99US-0131449P.
 PR 04-MAY-1999; 99US-0132048P.
 PR 05-MAY-1999; 99US-0132407P.
 PR 06-MAY-1999; 99US-0132484P.
 PR 07-MAY-1999; 99US-0132486P.
 PR 11-MAY-1999; 99US-0132863P.
 PR 14-MAY-1999; 99US-0134256P.
 PR 14-MAY-1999; 99US-0134218P.
 PR 14-MAY-1999; 99US-0134219P.
 PR 14-MAY-1999; 99US-0134221P.
 PR 18-MAY-1999; 99US-0134370P.
 PR 19-MAY-1999; 99US-0134768P.
 PR 20-MAY-1999; 99US-0134941P.
 PR 21-MAY-1999; 99US-0135124P.
 PR 24-MAY-1999; 99US-0135353P.
 PR 25-MAY-1999; 99US-0135629P.
 PR 27-MAY-1999; 99US-0136021P.
 PR 28-MAY-1999; 99US-0136392P.
 PR 01-JUN-1999; 99US-0136782P.
 PR 03-JUN-1999; 99US-0137222P.
 PR 04-JUN-1999; 99US-0137528P.
 PR 07-JUN-1999; 99US-0137502P.
 PR 08-JUN-1999; 99US-0137724P.
 PR 10-JUN-1999; 99US-0138094P.
 PR 99US-0138540P.

Db 2 AGCTCTTTGGACGTCTAAGCAAAACAAGATGTCGAGAGGCGTTTAGCGGTTACGAT 61
QY 85 AAGGACACACCCGACCGATGGCAATGTGGCAAAAGCTGTGGAGGGAACCTGTAGAA 144
Db 62 AAAGACACTCCGACCGTGGCAAAAGCTGTGAAGAGTTTGAAGTAATCTGCAGAG 121
QY 145 GAAGTGAAGCGCCACTATGACATTCCTGTCGAGGATCTCATCAACATCGAGACTGTCGT 204
Db 122 GAAGTTAAAGCGTCACTACGACATCCTCGTTGAAGATCTCATGAACATCGAACAAGACTTA 181
QY 205 GTCCCTTTGCCCAATTACAGACCTTCGA-----ATCTAACTCAAGAAGCATCAATGAC 258
Db 182 GTACCTTTGCTTAAATACAAAACCGTCGATGTTTGAAGTAATCTTAGAGGAATCAATGAT 241
QY 259 TTGACACAAAGGTATATACTAA 281
Db 242 TTGATTTGAGGTATTGAAGAA 264

RESULT 5

ID ADW16587 standard; cDNA; 482 BP.
XX
AC ADW16587;
XX
DT 24-MAR-2005 (first entry)
XX
DE Eucalyptus grandis transcription factor cDNA MYB family Seq 306.

XX gene; ss; plant; transcription; gene regulation; gene expression;
KW transgenic plant; drought resistance; disease resistance; salt tolerance;
KW cold tolerance; freezing tolerance; flowering; flavor enhancer;
KW flower color.
XX
OS Eucalyptus grandis.

XX
PN WO2005001050-A2.
XX
PD 06-JAN-2005.
XX
PF 07-JUN-2004; 2004WO-US017965.
XX
PR 06-JUN-2003; 2003US-0476189P.

XX (ARBO-) ARBORGEN LLC.
XX

PI Bloksberg LN, Bryant C, Connert MB, Emerson SJ, Frost MJ;
PI Forster RL, Grigor M, Higgins C, Lasham A, Lund ST, Magusin A;
PI Phillips J, Puthigae S, Veerakone S, Westwood C, Gause K, Wood M;
XX
DR WPI; 2005-075542/08.
DR P-PSDB; ADW17373.
XX
PT New polynucleotides isolated from plants encoding transcription factors,
PT and polypeptides encoded by such polynucleotides, useful for regulating
PT gene transcription and gene expression.
XX
PS Claim 3; SEQ ID NO 306; 1265pp; English.

XX This invention relates to novel isolated plant nucleic acid molecules, or
CC variants thereof, that encode transcription factors. Specifically, it
CC refers to transcription factor proteins that are capable of binding to
CC DNA in order to regulate gene transcription and gene expression in a
CC plants, in particular Eucalyptus grandis and Pinus radiata. The present
CC invention describes DNA constructs containing DNA encoding a
CC transcription factor that regulates the promoter, which is operably
CC linked to the desired nucleic acid to be expressed. It further provides
CC transgenic plants expressing a transcription factor that confers a trait
CC to the plant such as increased drought, salt or disease tolerance, height
CC change, enhanced cold/ frost tolerance, enhanced color, health and
CC nutritional characteristics, as well as improved taste, starch
CC composition, flower longevity and germination, amongst others.
CC Accordingly, such plants that are successfully transfected with a DNA

CC construct can be characterized by a difference in flower color, petal or
CC leaf shape and size, aroma or plant height. This polynucleotide is a
CC plant transcription factor cDNA sequence of the invention.

XX
SQ Sequence 482 BP; 137 A; 89 C; 121 G; 135 T; 0 U; 0 Other;

Query Match 32.2%; Score 122.8; DB 14; Length 482;
Best Local Similarity 70.9%; Pred. No. 8.5e-25;
Matches 163; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 28 TCACCATGACGTTTAGTCAAAACAAGATGTTTCGAGAGGCGCTTGCAGTTTACGACAAG 87

Db 64 TCATCTGACGCCGGAAGCAGAACAAAGATGTTTCGAGAAGGCATTTGCTCAATATGACAAG 123

QY 88 GACACACCCGACCGATGGCAACATGTGGCAAAAGCTGTGGAGGGAACCTGTAGAAGAA 147

Db 124 GACACTCCCGACAGGTGGCAGAAGATTGCCAAGGCCGTGGGGAATCTGCAGATGAA 183

QY 148 GTGAAGCGCCACTATGACATTTCTCGTCGAGGATCTCATCAACATCGAGACTGTCTGTC 207

Db 184 GTGAAGAAGACACTATGAATTTTAAATCGAGGACGTCAAGCACATCGAGTCTGCGAGATT 243

QY 208 CCTTTGCCCAATTACAAGACCTTCAATCTAACTCAAGAAGCATCAATGA 257

Db 244 CCTTTCTTAATTACAGGTTCGAGCAACAATAGCGGCTGAAGGTTCGGTGA 293

RESULT 6

ADR59489
ID ADR59489 standard; cDNA; 760 BP.

XX
AC ADR59489;

XX
DT 02-DEC-2004 (first entry)

XX
DE Cotton cDNA sequence, SEQ ID 270.

XX Cotton; ss; plant; cold tolerance; growth rate; cell cycle pathway;
KW drought tolerance; plant disease resistance; galactomannan; lignin;
KW plant growth regulator; heat tolerance; herbicide tolerance;
KW homologous recombination; extreme osmotic condition tolerance;
KW pathogen resistance; pest resistance; yield; photosynthesis; seed oil;
KW stress resistance.
XX
OS Gossypium hirsutum.

XX
PN US2004181830-A1.
XX
PD 16-SEP-2004.

XX
PF 29-JAN-2004; 2004US-00767795.
XX
PR 07-MAY-2001; 2001US-00849529.
PR 12-DEC-2001; 2001US-00021323.

XX
PA (KOVA/) KOVALIC D K.
PA (ZHOU/) ZHOU Y.
PA (CAO/) CAO Y.

XX
PI Kovalic DK, Zhou Y, Cao Y;
XX
DR WPI; 2004-667718/65.

XX New recombinant nucleic acid molecules and polypeptides from Gossypium
PT hirsutum, useful for producing plants with improved biological
*PT characteristics (e.g. improved plant cold or drought tolerance).
XX
PS Claim 1; SEQ ID NO 270; 14pp; English.

XX The invention relates to a recombinant polynucleotide comprising any of
CC the 58798 Cotton plant cDNA sequences mentioned in the specification.
CC Also a recombinant polypeptide comprising any of the 58798 amino acid
CC sequences mentioned in the specification and producing a plant having an

CC improved property. Producing a plant having an improved property
CC comprises transforming a plant with a recombinant construct comprising a
CC promoter region functional in a plant cell operably joined to a
CC polynucleotide comprising a coding sequence for a polypeptide associated
CC with the property, and growing the transformed plant. The polypeptide is
CC useful for improving plant cold tolerance, manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, improving plant
CC drought tolerance, providing increased resistance to plant disease,
CC producing galactomannan (or lignin or plant growth regulators), improving
CC plant heat tolerance, improving plant tolerance to herbicides, increasing
CC the rate of homologous recombination in plants, improving plant tolerance
CC to extreme osmotic conditions or to pathogens or pests, improving yield
CC by modification of photosynthesis, modifying seed oil or protein yield
CC and/or content, improving yield by modification of carbohydrate, nitrogen
CC or phosphorus use and/or uptake, or improving yield by providing improved
CC plant growth and development under at least one stress condition. The
CC polynucleotide and polypeptide may also be used in recombinant DNA
CC constructs, in physical arrays of molecules, as plant breeding markers,
CC or in computer-based storage and analysis systems. The present sequence
CC is a Cotton plant cDNA of the invention. NOTE: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20040181830. However only 6585
CC polynucleotide sequences were available, the remaining 52213
CC polynucleotides and all 58798 protein sequences were not present.

XX Sequence 760 BP; 308 A; 90 C; 144 G; 218 T; 0 U; 0 Other;

Query Match 31.7%; Score 120.6; DB 13; Length 760;
Best Local Similarity 75.4%; Pred. No. 4.3e-24;
Matches 150; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 32 CATGGACGTTTAGTCAAAACAAGATGTTCCGAGAGGGCCTTGGCAGTTTACGACAAGGACA 91
DB 138 CATGGACAGCCAGCAAAACAAGATTTCCGAAGGGCTTTAGCTGTTTACGACAAGGACA 197
QY 92 CACCCGACCGATGCGACAATGTGGCAAAAGCTGTCCGAGGGGAAACATGTAGAGAAGTGA 151
DB 198 CACGAGATCGTTGGTACAATGTTGCTAAAGCTGTGGGAGGGAACATGTTGAGGAAGTGA 257
QY 152 AGCGCCACTATGACATTTCTGTCGAGGATCTCATCAATCGAGACTGTCGTCGTCCTT 211
DB 258 AGAAGCACTATGAGCTTCTTCTGGAAGATGTTAGACACATCGAGTCGGGTCGGGTTCCCTT 317
QY 212 TGCCCAATTACAAGACCTT 230
DB 318 CCCCCGACTATTGGACCGT 336

RESULT 7

ACN61386/c

ID ACN61386 standard; cDNA; 543 BP.

XX ACN61386;

DT 02-DEC-2004 (first entry)

XX Cotton gynoecium tissue EST Clone ID: LIB3829-015-Q1-N6-E2, SEQ:16167.

XX Cotton; plant; EST; expressed sequence tag; transgenic plant; gynoecium;
KW variety Nucotton338; library LIB3829; molecular tag; molecular marker;
KW genetic mapping; molecular mapping; seed germination; plant growth;
KW plant quality; plant yield; plant breeding; tissue printing; ss.

XX Gossypium hirsutum.

XX US2004123340-A1.

XX 24-JUN-2004.

XX 12-DEC-2001; 2001US-00021323.

XX 14-DEC-2000; 2000US-0255619P.

XX (DEIK/) DEIKMAN J.
PA (FENG/) FENG P C C.
PA (FINC/) FINCHER K L.
PA (ZIEG/) ZIEGLER T E.
XX Deikman J, Feng PCC, Fincher KL, Ziegler TE;
PI WPI; 2004-479808/45.
XX New isolated nucleic acid molecule that encodes a plant protein or its
PT fragment, useful for isolating a variety of agronomically significant
PT genes associated with plant growth, quality or yield, and as molecular
PT tags to map genes.
XX Claim 1; SEQ ID NO 16167; 34pp; English.

XX The invention relates to 17880 cotton expressed sequence tags (ESTs;
CC ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated
CC from primed or non-primed seeds from variety DP50B, mature seeds from
CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoecium
CC tissue, developing fibres, carpel walls and septa from variety
CC Nucotton33B. The invention also relates to substantially purified
CC proteins or their fragments encoded by nucleic acid molecules of the
CC invention, and to transformed plants having a nucleic acid construct
CC comprising a nucleic acid of the invention. The cotton ESTs are useful as
CC molecular tags to isolate genetic regions, to isolate genes, to map
CC genes, to determine gene function and to determining whether genes are
CC members of a particular gene family. The nucleic acid molecules may be
CC used for isolating a variety of agronomically significant genes
CC associated with plant growth, quality, yield, and could also serve as
CC links in metabolic and catabolic pathways. The nucleic acid molecules are
CC also useful for identifying genes important in initiating and maintaining
CC seed germination or that may be used to mitigate stresses encountered
CC during seed germination. The ESTs additionally enable the acquisition of
CC promoters and cis-regulatory elements which will be useful to express
CC agronomically significant genes in these tissues and/or other tissues,
CC and also permits the acquisition of molecular markers useful in breeding
CC schemes, genetic and molecular mapping, and in cloning of agronomically
CC significant genes. The nucleic acid molecules are further useful for
CC detecting the expression level or pattern of a protein or mRNA and for
CC detecting the presence or quantity of a protein by tissue printing. The
CC present sequence represents a specifically claimed EST isolated from a
CC cotton variety Nucotton33B gynoecium tissue cDNA library (LIB3829). The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the US
CC patent office at seqdata.uspto.gov/sequence.html?DocID=US20040123340

XX Sequence 543 BP; 164 A; 128 C; 74 G; 177 T; 0 U; 0 Other;

Query Match 31.2%; Score 119; DB 13; Length 543;
Best Local Similarity 74.9%; Pred. No. 1.1e-23;
Matches 149; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 32 CATGGACGTTTAGTCAAAACAAGATGTTCCGAGAGGGCCTTGGCAGTTTACGACAAGGACA 91
DB 505 CATGGACAGCCAGCAAAACAAGATTTCCGAAGGGCTTTAGCTGTTTACGACAAGGACA 446
QY 92 CACCCGACCGATGCGACAATGTGGCAAAAGCTGTCCGAGGGGAAACATGTAGAGAAGTGA 151
DB 445 CACGAGATCGTTGGTACAATGTTCTTAAGCTGTGGAGAGAAACATGTTGAGGAAGTGA 386
QY 152 AGCGCCACTATGACATTTCTGTCGAGGATCTCATCAACATCGAGACTGTCGTCGTCCTT 211
DB 385 AGAAGCACTATGAGCTTCTTCTTGAAGATGTTAGACGATCGAGTCGGGTCGGGTTCCCTT 326
QY 212 TGCCCAATTACAAGACCTT 230
DB 325 TCCCCGACTATTGGACCGT 307

RESULT 8

ADR59490

ID ADR59490 standard; cDNA; 779 BP.
AC ADR59490;
XX
DT 02-DEC-2004 (first entry)
XX
DE Cotton cDNA sequence, SEQ ID 271.
XX
XX Cotton; ss; plant; cold tolerance; growth rate; cell cycle pathway;
KW drought tolerance; plant disease resistance; galactomannan; lignin;
KW plant growth regulator; heat tolerance; herbicide tolerance;
KW homologous recombination; extreme osmotic condition tolerance;
KW pathogen resistance; pest resistance; yield; photosynthesis; seed oil;
KW stress resistance.
XX
OS Gossypium hirsutum.
XX
PN US2004181830-A1.
XX
PD 16-SEP-2004.
XX
XX 29-JAN-2004; 2004US-00767795.
XX
PF 07-MAY-2001; 2001US-00849529.
XX
PR 12-DEC-2001; 2001US-00021323.
XX
PA (KOVA/) KOVALIC D K.
XX
PA (ZHOU/) ZHOU Y.
XX
XX (CAOY/) CAO Y.
PI Kovalic DK, Zhou Y, Cao Y;
XX
XX WPI; 2004-667718/65.
XX
XX New recombinant nucleic acid molecules and polypeptides from Gossypium
PT hirsutum, useful for producing plants with improved biological
PT characteristics (e.g. improved plant cold or drought tolerance).
XX
PS Claim 1; SEQ ID NO 271; 14pp; English.
XX
XX The invention relates to a recombinant polynucleotide comprising any of
CC the 58798 Cotton plant cDNA sequences mentioned in the specification.
CC Also a recombinant polypeptide comprising any of the 58798 amino acid
CC sequences mentioned in the specification and producing a plant having an
CC improved property. Producing a plant having an improved property
CC comprises transforming a plant with a recombinant construct comprising a
CC promoter region functional in a plant cell operably joined to a
CC polynucleotide comprising a coding sequence for a polypeptide associated
CC with the property, and growing the transformed plant. The polypeptide is
CC useful for improving plant cold tolerance, manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, improving plant
CC drought tolerance, providing increased resistance to plant disease,
CC producing galactomannan (or lignin or plant growth regulators), improving
CC plant heat tolerance, improving plant tolerance to herbicides, increasing
CC the rate of homologous recombination in plants, improving plant tolerance
CC to extreme osmotic conditions or to pathogens or pests, improving yield
CC by modification of photosynthesis, modifying seed oil or protein yield
CC and/or content, improving yield by modification of carbohydrate, nitrogen
CC or phosphorus use and/or uptake, or improving yield by providing improved
CC plant growth and development under at least one stress condition. The
CC polynucleotide and polypeptide may also be used in recombinant DNA
CC constructs, in physical arrays of molecules, as plant breeding markers,
CC or in computer-based storage and analysis systems. The present sequence
CC is a Cotton plant cDNA of the invention. NOTE: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequences.html?DocID=20040181830. However only 6585
CC polynucleotide sequences were available, the remaining 52213
CC polynucleotides and all 58798 protein sequences were not present.
SQ Sequence 779 BP; 248 A; 121 C; 157 G; 253 T; 0 U; 0 Other;
Query Match 31.2%; Score 119; DB 13; Length 779;

Best Local Similarity 74.9%; Pred. No. 1.3e-23;
Matches 149; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 32 CATGACGTTTGTCTCAAAACAAGATGTTTCGAGAGGCGCTTGGCAGTTTACGACAAGACA 91
DB 172 CATGGACAGCCCAAGCAAAAACAAAGATTTTCGAAAGGCTTTAGCTGTTTACGACAAGACA 231
QY 92 CACCCGACCGATGGCACAATGTGGCAAAAGCTGTGCGAGGGAACCTGTAGAGAAGTGA 151
DB 232 CACCAAGATCGTTGGTACATGTTGCTTAAGCTGTGGAGAGAAAACCTGTTGAGGAAGTGA 291
QY 152 AGCGGCACCTATGACATTTCTTCGAGGATCTCATCAACATCGAGACTGGTGTGTCCTT 211
DB 292 AGAAGCACTATGAGCTTCTTCTTGAAGATGTTAGAGCATCGAGTCGGTTCGGTTTCTT 351
QY 212 TCGCCCAATTACAAGACCTT 230
DB 352 TCCCGACTATTGGACCGT 370
RESULT 9
ADX32012
ID ADX32012 standard; cDNA; 704 BP.
XX
AC ADX32012;
XX
DT 21-APR-2005 (first entry)
XX
XX Plant full length insert polynucleotide seqid 14832.
XX
KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content; gene; ss.
XX
XX Unidentified.
XX
XX US2004034888-A1.
PN
PD 19-FEB-2004.
XX
XX 28-APR-2003; 2003US-00425114.
PF
XX 06-MAY-1999; 99US-00304517.
PR
PR 05-NOV-2001; 2001US-00985678.
XX
XX (LIUJ/) LIU J.
XX (ZHOU/) ZHOU Y.
XX (KOVA/) KOVALIC D K.
XX (SCRE/) SCREEN S E.
XX (TABA/) TABASKA J E.
XX (CAOY/) CAO Y.
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX
XX WPI; 2004-180133/17.
XX
XX New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
XX Claim 1; SEQ ID NO 14832; 15pp; English.
PS
XX The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.seqdata.uspto.gov/sequence.html?DocID=2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as

CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This sequence represents a plant full length insert
 CC polynucleotide that can be used in the recombinant DNA construct of the
 CC invention.

SQ Sequence 704 BP; 202 A; 159 C; 121 G; 222 T; 0 U; 0 Other;

Query Match 29.1%; Score 110.8; DB 13; Length 704;
 Best Local Similarity 68.1%; Pred. No. 2.7e-21;
 Matches 154; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
 QY 1 ATGGCGTCAAACTCAAGAGTTCAATCTCACCATTGGACGTTTGTAGTCAAAACAAGATGTTTC 60
 DB 150 ATGGATCCAGTTCATCTCAGCTCTGGCTCATGGAGTTAAGGACACACAGGCGCTTT 209
 QY 61 GAGAGGCGCTTGGCAGTTTACGACAGGACACACCGACCGATGGCACATGTGGCAAAA 120
 DB 210 GAAAAGCGCTTTAGCTGTTTATGACAGGACACTCTGACCGTTGGTACAATGTTGCTCAT 269
 QY 121 GCTGTGGAGGAAAACGTGAGAAGAGTGAAGGCCCACTATGACATTCCTCGCAGGAT 180
 DB 270 GCTGTTGGTGGCAAACTCCAGAGGAAGTGAAGGCACTACGAACTCTTGTTCAGGAT 329
 QY 181 CTATCAACATCGAGACTGTGTCCTCTTGGCCCAATTACAAGA 226
 DB 330 GTTAAGCATATTGAGTCTGGACGTGTCCTATTCCTCAAAATTACAGA 375

RESULT 10
 ADJ43786/c
 ID ADJ43786 standard; cDNA; 466 BP.

XX ADJ43786;

XX 06-MAY-2004 (first entry)

XX Plant cDNA #4786.

XX Plant; gene; ss; transcription; plant genome augmentation; cereal;
 KW soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;
 KW maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;
 KW stress tolerance; salt tolerance; cold tolerance; drought tolerance;
 KW plant nutrition; apical dominance; dwarfism; early flowering; antiviral;
 KW antifungal.

XX Eukaryota.

XX US2004016025-A1.

XX 22-JAN-2004.

XX 26-SEP-2002; 2002US-00260238.

XX 26-SEP-2001; 2001US-0325277P.

XX 26-SEP-2001; 2001US-0325448P.

XX 04-APR-2002; 2002US-0370620P.

XX (BUDW/) BUDWORTH P.

XX (MOUG/) MOUGHAMER T.

XX (BRIG/) BRIGGS S P.

XX (COOP/) COOPER B.

XX (GLAZ/) GLAZEBROOK J.

XX (GOFF/) GOFF S A.

XX (KATA/) KATAGIRI F.

PA (KREP/) KREPS J.
 PA (PROV/) PROVART N.
 PA (RICK/) RICHE D.
 PA (ZHUT/) ZHU T.
 XX Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J;
 PI Goff SA, Katagiri F, Kreps J, Provart N, Riche D, Zhu T;
 XX WPI; 2004-190374/18.

XX New rice promoter, useful for manipulating crop plants to alter or
 PT improve phenotypic characteristics, e.g. produce large quantities of oil
 PT or proteins, resistance to insecticides, virus or fungi, stress tolerance
 PT or high nutritional value.

XX Example 13; SEQ ID NO 4786; 230pp; English.

XX The invention relates to plant nucleotide sequences that direct seed-,
 CC leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential
 CC or constitutive transcription of an operatively linked nucleic acid
 CC segment. The invention also relates to a method for augmenting a plant
 CC genome and a method of identifying a gene, where its expression is
 CC altered in the seed, leaf, stem, panicle, pollen, root or is constitutive
 CC in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,
 CC canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,
 CC sorghum, rice or wheat. The polynucleotides and the polypeptides they
 CC encode are useful for manipulating crop plants to alter or improve
 CC phenotypic characteristics, to produce large quantities of oil or
 CC proteins, to incur resistance to insecticides, viruses or fungi, and to
 CC incur stress tolerance (e.g. salt, cold or drought) to ensure the plants
 CC have a high nutritional value with reduced apical dominance or dwarfism,
 CC early flowering or altered metabolic pathways. This sequence represents a
 CC plant nucleic acid of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification but was obtained in
 CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 466 BP; 93 A; 127 C; 124 G; 122 T; 0 U; 0 Other;

Query Match 28.9%; Score 110; DB 12; Length 466;
 Best Local Similarity 67.4%; Pred. No. 3.8e-21;
 Matches 155; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 25 ATCTCACCATGGACGTTTGTAGTCAAAACAAGATTTTCGAGAGGCGCTTGGCAGTTTACGAC 84

DB 407 AGCTCGTGTGGACCGCGGAGCAGACAAAGATTTTCGAGAAAGCCCTCGCGGTGTACGAC 349

QY 85 AAGGACACACCCGACCGATGGCACAATGTGGCAAAAGCTGTCGAGAGGAAAACCTGTAGAA 144

DB 347 AAGGACACGCGCCGACCGCTGGCACACAGTGGCGCGCGCGCTCGCGGCAAGTCAGCGGAA 288

QY 145 GAAAGTGAAGCGCCACTATGACATTTCTCGTCGAGGATCTCATCAACATCGAGACTGTGCGT 204

DB 287 GAAGTGAAGCGGCACCTACGAGCTGCTGTTGGCGGACATTATCTCTCATCGAGAGGGCCAG 228

QY 205 GTCCTTTGCCCAATTACAAAGACCTTCGAATCTTAACCTCAAGACATCAA 254

DB 227 ATGCTCGAGCCAATTACCGCTCTCCGGCCACAGGGGATGAAGTGTCGA 178

RESULT 11

ACL30919/c

ID ACL30919 standard; cDNA; 513 BP.

XX ACL30919;

XX 02-JUN-2005 (first entry)

XX Rice abiotic stress response related polynucleotide SEQ ID NO:9482.

XX ss; abiotic stress tolerance; transgenic plant; plant; cereal;

XX agriculture.

XX Oryza sativa.

OS

CC also useful for identifying genes important in initiating and maintaining
CC seed germination or that may be used to mitigate stresses encountered
CC during seed germination. The ESTs additionally enable the acquisition of
CC promoters and cis-regulatory elements which will be useful to express
CC agronomically significant genes in these tissues and/or other tissues,
CC and also permits the acquisition of molecular markers useful in breeding
CC schemes, genetic and molecular mapping, and in cloning of agronomically
CC significant genes. The nucleic acid molecules are further useful for
CC detecting the expression level or pattern of a protein or mRNA and for
CC detecting the presence or quantity of a protein by tissue printing. The
CC present sequence represents a specifically claimed EST isolated from a
CC cotton variety Nu cotton33B gynoecium tissue cDNA library (LIB3829). The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the US
CC patent office at seqdata.uspto.gov/sequence.html?DocID=US20040123340
XX

SQ Sequence 505 BP; 131 A; 114 C; 102 G; 158 T; 0 U; 0 Other;

Query Match 28.2%; Score 107.4; DB 13; Length 505;
Best Local Similarity 70.2%; Pred. No. 2.2e-20;
Matches 144; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 22 TCAATCTCACCATGGAGCTTTAGTCAAAACAAGATGTCGAGAGGCGCTTGGCAGTTTAC 81
Db 466 TCAGGTTGCAACTGGACGCCAAGCAAAACAAGTTGTCGAGAACGCTTTAGTATCTAC 407
Qy 82 GACAAGGACACACCCGACCGCATGCAATGTGGCAAAAGCTGTGCGAGGGGAAACTGTA 141
Db 406 GACAAGGATACACGACCGGTGGCACAAGCTAGCCAGGCGTGTGGGGGAGACCGTG 347
Qy 142 GAAGAAGTGAAGCGCCACTATGACATTCCTCGTGAGGATCTCATCAACATCGAGACTGGT 201
Db 346 GAGGAAGTGAAGTTGATTACCAAGACCTTGTGGATGACATCAAGCAGATAGAGTCTGGG 287
Qy 202 CGTGTCCTTTGCCAATTACAAGA 226
Db 286 CACGTGCTCTTTGCCCTTTACAAGA 262

RESULT 15
ACN59701
ID ACN59701 standard; cDNA; 532 BP.
XX
AC ACN59701;
XX
DT 02-DEC-2004 (first entry)
XX
XX Cotton gynoecium tissue EST Clone ID: LIB3829-034-Q1-K6-A10, SEQ:14482.
DE

XX Cotton; plant; EST; expressed sequence tag; transgenic plant; gynoecium;
KW variety Nu cotton33B; library LIB3829; molecular tag; molecular marker;
KW genetic mapping; molecular mapping; seed germination; plant growth;
KW plant quality; plant yield; plant breeding; tissue printing; ss.

OS Gossypium hirsutum.
XX
XX US2004123340-A1.
XX
XX 24-JUN-2004.
XX

PF 12-DEC-2001; 2001US-00021323.
XX
XX 14-DEC-2000; 2000US-0255619P.
XX

PA (DEIK/) DEIKMAN J.
PA (FENG/) FENG P C C.
PA (FING/) FINCHER K L.
PA (ZIEG/) ZIEGLER T E.

XX Deikman J, Feng PCC, Fincher KL, Ziegler TE;
PI
XX WPI; 2004-479808/45.
XX

PT New isolated nucleic acid molecule that encodes a plant protein or its
PT fragment, useful for isolating a variety of agronomically significant
PT genes associated with plant growth, quality or yield, and as molecular
PT tags to map genes.
XX

PS Claim 1; SEQ ID NO 14482; 34pp; English.
XX
CC The invention relates to 17880 cotton expressed sequence tags (ESTs;
CC ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated
CC from primed or non-primed seeds from variety DP50B, mature seeds from
CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoecium
CC tissue, developing fibres, carpel walls and septa from variety
CC Nu cotton33B. The invention also relates to substantially purified
CC proteins or their fragments encoded by nucleic acid molecules of the
CC invention, and to transformed plants having a nucleic acid construct
CC comprising a nucleic acid of the invention. The cotton ESTs are useful as
CC molecular tags to isolate genetic regions, to isolate genes, to map
CC genes, to determine gene function and to determining whether genes are
CC members of a particular gene family. The nucleic acid molecules may be
CC used for isolating a variety of agronomically significant genes
CC associated with plant growth, quality, yield, and could also serve as
CC links in metabolic and catabolic pathways. The nucleic acid molecules are
CC also useful for identifying genes important in initiating and maintaining
CC seed germination or that may be used to mitigate stresses encountered
CC during seed germination. The ESTs additionally enable the acquisition of
CC promoters and cis-regulatory elements which will be useful to express
CC agronomically significant genes in these tissues and/or other tissues,
CC and also permits the acquisition of molecular markers useful in breeding
CC schemes, genetic and molecular mapping, and in cloning of agronomically
CC significant genes. The nucleic acid molecules are further useful for
CC detecting the expression level or pattern of a protein or mRNA and for
CC detecting the presence or quantity of a protein by tissue printing. The
CC present sequence represents a specifically claimed EST isolated from a
CC cotton variety Nu cotton33B gynoecium tissue cDNA library (LIB3829). The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the US
CC patent office at seqdata.uspto.gov/sequence.html?DocID=US20040123340
XX

SQ Sequence 532 BP; 170 A; 104 C; 119 G; 139 T; 0 U; 0 Other;
Query Match 28.2%; Score 107.4; DB 13; Length 532;
Best Local Similarity 70.2%; Pred. No. 2.2e-20;
Matches 144; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 22 TCAATCTCACCATGGAGCTTTAGTCAAAACAAGATGTCGAGAGGCGCTTGGCAGTTTAC 81
Db 56 TCAGGTTGCAACTGGACGCCAAGCAAAACAAGTTGTCGAGAACGCTTTAGTATCTAC 115
Qy 82 GACAAGGACACACCCGACCGCATGCAATGTGGCAAAAGCTGTGCGAGGGGAAACTGTA 141
Db 116 GACAAGGATACACGACCGGTGGCACAAGCTAGCCAGGCGTGTGGGGGAGACCGTG 175
Qy 142 GAAGAAGTGAAGCGCCACTATGACATTCCTCGTGAGGATCTCATCAACATCGAGACTGGT 201
Db 176 GAGGAAGTGAAGTTGATTACCAAGACCTTGTGGATGACATCAAGCAGATAGAGTCTGGG 235
Qy 202 CGTGTCCTTTGCCAATTACAAGA 226
Db 236 CACGTGCTCTTTGCCCTTTACAAGA 260

Search completed: March 10, 2006, 18:43:29
Job time : 404 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 19:35:45 ; Search time 442 Seconds
(without alignments)
657.736 Million cell updates/sec

Title: US-10-697-787-2

Perfect score: 656

Sequence: 1 MASNGRSSIPMTFSQKMF.....HSSDFEKFQKLVSYISLV 126

Scoring table: BLOSUM62

Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 7673375 seqs, 1153648444 residues

Total number of hits satisfying chosen parameters: 15346750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abs/ABSSWEB_spool/US10697787/runat_09032006_150641_3661/app_query.fasta 1
-DB=Published Applications NA New -QFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs05p
-USER=US10697787 @CGN_1_1_653 @runat_09032006_150641_3661 -NCPUP=6 -ICPU=3
-NO_WMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA New.*

1:	/cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
2:	/cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3:	/cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4:	/cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
5:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
6:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
7:	/cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
8:	/cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
9:	/cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
10:	/cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
11:	/cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
12:	/cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
13:	/cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	269.5	41.1	517	9	US-11-096-568A-8355
2	214	32.6	1724	9	Sequence 8355, Ap
3	211	32.2	866	7	US-11-096-568A-4492
4	204	31.1	896	7	Sequence 4492, Ap
					Sequence 245, App
					Sequence 273, App

5	193	29.4	957	9	US-11-087-099-102	Sequence 102, App
6	187	28.5	876	9	US-11-096-568A-29719	Sequence 29719, A
7	185.5	28.3	1254	9	US-11-096-568A-21266	Sequence 21266, A
8	151.5	23.1	994	9	US-11-096-568A-23894	Sequence 23894, A
9	146	22.3	1351	9	US-11-096-568A-20386	Sequence 20386, A
10	126	19.2	1431	9	US-11-096-568A-13836	Sequence 13836, A
11	124	18.9	933	9	US-11-096-568A-24021	Sequence 24021, A
12	121	18.4	600	12	US-11-136-527-4714	Sequence 4714, App
13	121	18.4	2119	12	US-11-136-527-618	Sequence 618, App
C 14	96	14.6	1593	8	US-10-750-185-30681	Sequence 30681, A
C 15	96	14.6	1593	8	US-10-750-623-30681	Sequence 30681, A
C 16	83.5	12.7	607	6	US-09-925-065A-206387	Sequence 206387, A
C 17	83.5	12.7	607	6	US-09-925-065A-206388	Sequence 206388, A
18	83.5	12.7	2919	7	US-10-932-182A-3711	Sequence 3711, Ap
19	83.5	12.7	2919	7	US-10-932-182A-3711	Sequence 3711, Ap
20	82	12.5	1506	9	US-11-096-568A-22860	Sequence 22860, A
C 21	80.5	12.3	607	6	US-09-925-065A-206386	Sequence 206386, A
C 22	79	12.0	499	9	US-11-096-568A-14174	Sequence 14174, A
C 23	78.5	12.0	607	6	US-09-925-065A-140607	Sequence 140607, A
C 24	78.5	12.0	162173	12	US-11-121-086-72	Sequence 72, Appl
25	75.5	11.5	1509	9	US-11-096-568A-26086	Sequence 26086, A
26	75.5	11.5	1545	9	US-11-096-568A-9922	Sequence 9922, Ap
27	75	11.4	1332	9	US-11-096-568A-33483	Sequence 33483, A
28	75	11.4	3276	7	US-10-932-182A-190881	Sequence 190881, A
29	75	11.4	3276	7	US-10-932-182A-190881	Sequence 190881, A
30	74.5	11.4	1779	7	US-10-932-182A-79555	Sequence 79555, A
31	74.5	11.4	1779	7	US-10-932-182A-79555	Sequence 79555, A
32	74	11.3	885	7	US-10-932-182A-78621	Sequence 78621, A
33	74	11.3	885	7	US-10-932-182A-78621	Sequence 78621, A
34	74	11.3	2550	7	US-10-932-182A-77658	Sequence 77658, A
35	74	11.3	2550	7	US-10-932-182A-77658	Sequence 77658, A
36	73	11.1	864	9	US-11-096-568A-28818	Sequence 28818, A
37	72.5	11.1	1230	7	US-10-932-182A-82655	Sequence 82655, A
38	72.5	11.1	1230	7	US-10-932-182A-82655	Sequence 82655, A
39	71.5	10.9	2295	12	US-11-055-557-19	Sequence 19, Appl
40	71.5	10.9	2307	12	US-11-055-557-3	Sequence 3, Appl
41	71.5	10.9	2307	12	US-11-055-557-5	Sequence 5, Appl
42	71.5	10.9	6425	12	US-11-131-479-90	Sequence 90, Appl
43	71	10.8	1638	7	US-10-932-182A-79707	Sequence 79707, A
44	71	10.8	1638	7	US-10-932-182A-79707	Sequence 79707, A
45	70.5	10.7	1390	9	US-11-096-568A-5907	Sequence 5907, Ap

ALIGNMENTS

RESULT 1

US-11-096-568A-8355
; Sequence 8355, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096, 568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 8355
; LENGTH: 517
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(517)
; OTHER INFORMATION: Ceres Seq. ID no. 15225172
; US-11-096-568A-8355

Alignment Scores:	6.99e-25	Length:	517
Pred. No.:	269.50	Matches:	55
Score:	66.7%	Conservative:	13
Best Local Similarity:	53.9%	Mismatches:	31
Query Match:	41.1%	Indels:	3

```
DB: 9 Gaps: 1
US-10-697-787-2 (1-126) x US-11-096-568A-8355 (1-517)
Qy 3 SerAsnSerArgSerSerIleSerProTyrThrPheSerGlnAsnLysMetPheGluArg 22
Db 63 GCTACAAACATGGCCCTCAAGTCAGGTGGCTCCGAAGCAAGCAAGATTTGAGAT 122
Qy 23 AlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnValAlaLysAlaVal 42
Db 123 GCCTTGGCATCTTCGACAGGACACCCACAGACAGGTGGCACAGGTGGCCAGGCGCTC 182
Qy 43 GlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAspLeuIle 62
Db 183 GGAGGAAAAACGGTGGAGGAAGTGAAGGCATATTATGAGAAGCTCGTGGAAAGATGTAAG 242
Qy 63 AsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGluSerAsnSerArg 82
Db 243 GAGATTGAGGAAGGTACAGTGCCTCCCAATTACCGA-----AGTGTGCAAGA 293
Qy 83 SerIleAsnAspPheAspThrArgTyrIleThrLysTyrLeuTyrMetMetLeuSerIle 102
Db 294 GGCACCGTTACATGGATGAAGAAACAGGATGAGGTTCTGAGTCTGCAGTGAAGTGTG 353
Qy 103 TyrPhe 104
Db 354 ATTTAT 359
RESULT 2
US-11-096-568A-4492
; Sequence 4492, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 4492
; LENGTH: 1724
; TYPE: DNA
; ORGANISM: Glycine max
; NAME/KEY: misc feature
; LOCATION: (1)-(1724)
; OTHER INFORMATION: Ceres Seq. ID no. 13638441
US-11-096-568A-4492
Alignment Scores:
Pred. No.: 4,54e-17 Length: 1724
Score: 214.00 Matches: 41
Percent Similarity: 71.8% Conservative: 12
Best Local Similarity: 55.4% Mismatches: 19
Query Match: 32.6% Indels: 2
DB: 9 Gaps: 1
US-10-697-787-2 (1-126) x US-11-096-568A-4492 (1-1724)
Qy 3 SerAsnSerArgSerSerIleSerProTyrThrPheSerGlnAsnLysMetPheGluArg 22
Db 560 AATAATAGCAGGAGC-----ACAAAATGGACCTCTCAAGAGAGACAAAGCTCTTTGAAAT 613
Qy 23 AlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnValAlaLysAlaVal 42
Db 614 GCTCTTGAGTGCATGATAGACACCCCGGATCGGTGGCACAGTGGCTGAGATGATT 673
Qy 43 GlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAspLeuIle 62
Db 674 CCTGGAAGACAGTGGTTGATGTGTAAGGCAGTACAAGGAGTTGGAAGTAGATGTTAGC 733
Qy 63 AsnIleGluThrGlyArgValProLeuProAsnTyrLysThr 76
```

```
Db 734 AATATAGAAAGCTGGTTTGATCCAGTCTCGCTATAGTAGC 775
RESULT 3
US-10-714-887-245
; Sequence 245, Application US/10714887
; Publication No. US20060015972A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: HEARD, Jacqueline
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: CREELMAN, Robert
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: CANALES, Roger
; APPLICANT: REPETTI, Peter
; APPLICANT: KUMIMOTO, Roderick W
; APPLICANT: GUTTERSON, Neal
; APPLICANT: REUBER, T. Lynne
; APPLICANT: PINEDA, Omaika
; APPLICANT: SHERMAN, Bradley K
; TITLE OF INVENTION: PLANT TRANSCRIPTIONAL REGULATORS OF DROUGHT STRESS
; FILE REFERENCE: MBI0058-CIP
; CURRENT APPLICATION NUMBER: US/10/714,887
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: 10/412,699
; PRIOR FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/135,134
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 245
; LENGTH: 866
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G2701 reference sequence; predicted polypeptide sequence is para
US-10-714-887-245
Alignment Scores:
Pred. No.: 4,6e-17 Length: 866
Score: 211.00 Matches: 41
Percent Similarity: 64.6% Conservative: 10
Best Local Similarity: 51.9% Mismatches: 28
Query Match: 32.2% Indels: 0
DB: 7 Gaps: 0
US-10-697-787-2 (1-126) x US-10-714-887-245 (1-866)
Qy 1 MetAlaSerAsnSerArgSerSerIleSerProTyrThrPheSerGlnAsnLysMetPhe 20
Db 109 ATGGTAGCTTACACAGCTCGAGTACGGTAGCTGAGCTAAAGAGAGACAAGATGTC 168
Qy 21 GluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnValAlaLys 40
Db 169 GAACGAGCTCTTCGATATACGCTGAAGACTCGCTGATCGCTGGTTTAAAGTTGCTTCC 228
```

QY 41 AlaValGlyGlyLysThrValGluGluValLysArgHisThrAspLeuValGluAsp 60
Db 229 ATGATCCTGGAAGAGCTGTTTGTATGATGAAGCAATATAGTAAGCTTTGAAGAGAC 288
QY 61 LeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGluSer 79
Db 289 GTTTCGATATTGAAGAGAGAGCTGTGCCATTCTCGGTATCTCTCAGCTTCTCT 345

RESULT 4

US-10-714-887-273
; Sequence 273, Application US/10714887
; Publication No. US20060015972A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: HEARD, Jacqueline
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: CREELMAN, Robert
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: CANALES, Roger
; APPLICANT: REPETTI, Peter
; APPLICANT: KUMIMOTO, Roderick W
; APPLICANT: GUTTERSON, Neal
; APPLICANT: REUBER, T. Lynne
; APPLICANT: FINEDA, Omaira
; APPLICANT: SHERMAN, Bradley K
; TITLE OF INVENTION: PLANT TRANSCRIPTIONAL REGULATORS OF DROUGHT STRESS
; FILE REFERENCE: MB10058-CIP
; CURRENT APPLICATION NUMBER: US/10/714,887
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: 10/412,699
; PRIOR FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/135,134
; PRIOR FILING DATE: 1993-05-20
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1993-03-23
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 430
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 273
; LENGTH: 896
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:

US-10-714-887-273
; OTHER INFORMATION: G1634; predicted polypeptide sequence is paralogous to G2701

Alignment Scores:
Pred. No.: 3,83e-16 Length: 896
Score: 204.00 Matches: 38
Percent Similarity: 68.7% Conservative: 8
Best Local Similarity: 56.7% Mismatches: 21
Query Match: 31.1% Indels: 0
DB: 7 Gaps: 0

US-10-697-787-2 (1-126) x US-10-714-887-273 (1-896)

QY 10 SerProTrrPheSerGlnAsnLysMetPheGluArgAlaLeuAlaValTyrAspLys 29
Db 106 TCGAGCTGGACTTAAGAAGAGCAAGAGTTTGACGAGCTCTTGTCTACGCTGAT 165

QY 30 AspThrProAspArgTrrPheHisValAlaLysAlaValGlyGlyLysThrValGluGlu 49
Db 166 GACAGCCCTGATCGCTGGTTCAAGTTGCTGTATGATCCCTGGAAAGACCATATCAGAT 225
QY 50 ValLysArgHisTyrAspIleLeuValGluAspLeuIleAsnIleGluThrGlyArgVal 69
Db 226 GTCATGAGGCAATACTCTAAGCTTGAAGAAGACCTCTTCGATATCGAAGCAGGACTTGT 285
QY 70 ProLeuProAsnTyrLysThr 76
Db 286 CCGATCCCGGGTTACCGTTCA 306

RESULT 5

US-11-087-099-102
; Sequence 102, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 102
; LENGTH: 957
; TYPE: DNA
; ORGANISM: Glycine max
; US-11-087-099-102

Alignment Scores:

Pred. No.: 1,09e-14 Length: 957
Score: 193.00 Matches: 43
Percent Similarity: 48.5% Conservative: 7
Best Local Similarity: 41.7% Mismatches: 25
Query Match: 29.4% Indels: 28
DB: 9 Gaps: 2

US-10-697-787-2 (1-126) x US-11-087-099-102 (1-957)

QY 12 TrpThrPheSerGlnAsnLysMetPheGluArgAlaLeuAlaValTyrAspLysAspThr 31
Db 109 TGGACTAGAGAAGATAACAAGAGTTTGAAGTGCCCTTGCTATATATGACAAGGACACC 168
QY 32 ProAspArgTrrPheHisValAlaLysAlaValGlyGlyLysThrValGluGluValLys 51
Db 169 CCAGATAGATGTTCAAGGTGGCCGCTATGATCCCTGGGAAGACTGTGTTTGATGTGATC 228
QY 52 ArgHisTyrAspIleLeuValGluAspLeuIleAsnIleGluThrGlyArgValProLeu 71
Db 229 AAGCAATATAGGAAGACTCGAAGAGAGATGTGAGTGAATCGAAGCAGGCGCATGTTCCGATT 288
QY 72 ProAsnTyrLys-----ThrPheGluSerAsnSerArgSerIleAsnAspPhe 87
Db 289 CCGGCTACCTTGATCTCTTTCACCTTTCAG----- 321
QY 88 AspThrArgTyrIleThrLysTyrLeuTyrMetMetLeuSerIleTyrPheAsnAsnHis 107
Db 322 -----CTGTTTGACAACACCAC 336

QY 108 SerSerAsp 110

Db 337 AACTATGAT 345

RESULT 6

US-11-096-568A-29719
; Sequence 29719, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptide
; FILE REFERENCE: 2750-1592PUS2

```
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 29719
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)-(876)
; OTHER INFORMATION: Ceres Seq. ID no. 4926979
US-11-096-568A-29719

Alignment Scores:
Pred. No.: 5,76e-14 Length: 876
Score: 187.00 Matches: 35
Percent Similarity: 68.3% Conservative: 8
Best Local Similarity: 55.6% Mismatches: 20
Query Match: 28.5% Indels: 0
DB: 9 Gaps: 0

US-10-697-787-2 (1-126) x US-11-096-568A-29719 (1-876)

Qy 12 TrpThrPheSerGlnAsnLysMetPheGluArgAlaLeuAlaValTyrAspLysAspThr 31
Db 64 TGGACAGCTGAAGAAACAAAGAAATTCGAAACGCITTTAGCCTTTTACGACAAAGATACT 123

Qy 32 ProAspArgTrpHisAsnValAlaLysAlaValGlyGlyLysThrValGluGluValLys 51
Db 124 CCCGACAGATGGTCCAGAGTCGTCGCCATCTCCCGGTAAACAGTCGGAGATGTGATC 183

Qy 52 ArgHisTyrAspLeuValGluAspLeuLeuAsnIleGluThrGlyArgValProLeu 71
Db 184 AAACAATACAGAGAGCTTGAGGAAGACGTTAGGCACATCGAAGCTGTCTTATACCAATC 243

Qy 72 ProAsnTyr 74
Db 244 CTTGGTTAC 252

RESULT 7
US-11-096-568A-21266
; Sequence 21266, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 21266
; LENGTH: 1254
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)-(1254)
; OTHER INFORMATION: Ceres Seq. ID no. 12402516
US-11-096-568A-21266

Alignment Scores:
Pred. No.: 1.42e-13 Length: 1254
Score: 185.50 Matches: 41
Percent Similarity: 61.4% Conservative: 10
Best Local Similarity: 49.4% Mismatches: 23
Query Match: 28.3% Indels: 2
DB: 9 Gaps: 2

US-10-697-787-2 (1-126) x US-11-096-568A-21266 (1-1254)

Qy 1 MetAlaSerAsnSerArgSerSer-----IleSerProTyrThr 13
||||| ||| |||||
```

```
Db 151 ATGGCCGTCGAACGAGCGGAGCAGCAGTCGCGGTGGGGAGGCGGTGGTGGTGGTGGACC 210
Qy 14 PheSerGlnAsnLysMetPheGluArgAlaLeuAlaValTyr-----AspLysAspThr 31
||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 211 CGCGAGCAGGAGAGCGGTTTCGAGAACCGCGTTCGCGACATGGCGCGGAGGAGGACGGG 270
Qy 32 ProAspArgTrpHisAsnValAlaLysAlaValGlyGlyLysThrValGluGluValLys 51
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 271 GACGCGAGGTGGGAGAGTCGCGAGGCCGTCGAGGGGAAGACGCCGCGGAGGAGGTACGG 330
Qy 52 ArgHisTyrAspLeuValGluAspLeuLeuAsnIleGluThrGlyArgValProLeu 71
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 331 CGGCACTACGAGCTGCTGCTGGAGGACGTCGACGCGCATCGAGTCGGCGCGGCTCCCGCTT 390

Qy 72 ProAsnTyr 74
||| |||
Db 391 CCGGCGGTAC 399

RESULT 8
US-11-096-568A-23894
; Sequence 23894, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 23894
; LENGTH: 994
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)-(994)
; OTHER INFORMATION: Ceres Seq. ID no. 12416292
US-11-096-568A-23894

Alignment Scores:
Pred. No.: 2.52e-09 Length: 994
Score: 151.50 Matches: 30
Percent Similarity: 64.1% Conservative: 11
Best Local Similarity: 46.9% Mismatches: 22
Query Match: 23.1% Indels: 1
DB: 9 Gaps: 1

US-10-697-787-2 (1-126) x US-11-096-568A-23894 (1-994)

Qy 12 TrpThrPheSerGlnAsnLysMetPheGluArgAlaLeuAlaValTyrAspLysAspThr 31
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 204 TGGACGCGGAGGAGAACAAAGCTGTTTCGAGAAAGGCACTGGCGCAGATCGCACCGAAGCGG 263

Qy 32 ProAspArgTrpHisAsnValAlaLysAlaValGlyGlyLysThrValGluGluValLys 51
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 264 CCCGACAGGTGGGAGAGGTGGCGCGGTGCTCCCTCGGAGAGCGGTTCGACGACGTGAGG 323

Qy 52 ArgHisTyrAspLeuValGluAspLeuLeuAsnIleGluThr---GlyArgValPro 70
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 324 AGCCAGTACCACGCCCTCGAGAGAGGACGTGGGGCTCATCGAGGCGCGGCGCTCGTCCG 383

Qy 71 LeuProAsnTyr 74
||| |||
Db 384 TTCCCGCGGTAC 395

RESULT 9
US-11-096-568A-20386
; Sequence 20386, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 19:35:01 ; Search time 657 Seconds
(without alignments)
1585.908 Million cell updates/sec

Title: US-10-697-787-2

Perfect score: 656

Sequence: 1 MASNRGRSSIPWTFQNKMF.....HSSDFEKFQKLVSVISIV 126

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abs9/ABSSWEB_spool/US10697787/runat_09032006_150638_3580/app_query.fasta_1
-DB=Published_Applications NA_Main -QFWT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human4.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pco -NORM=ext
-HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs04
-USER=US10697787 @CGN_1_1_1026 @runat_09032006_150638_3580 -NCPU=6 -ICPU=3
-NO_WMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA_Main:

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq*
- 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	656	100.0	381	8	US-10-697-787-1		Sequence 1, Appli
2	656	100.0	381	9	US-10-512-600-1		Sequence 1, Appli
3	293	44.7	678	7	US-10-424-599-6644		Sequence 6644, Ap
4	290	44.2	760	8	US-10-767-795-270		Sequence 270, App
5	287	43.8	704	7	US-10-425-114-14832		Sequence 14832, A
6	287	43.8	730	7	US-10-424-599-135729		Sequence 135729,
7	286	43.6	556	3	US-09-770-152-473		Sequence 473, App

8	286	43.6	578	6	US-10-225-066A-359	Sequence 359, App
9	286	43.6	578	7	US-10-374-780A-2719	Sequence 2719, Ap
10	286	43.6	578	9	US-10-225-068A-359	Sequence 359, App
11	278	42.4	732	8	US-10-697-787-18	Sequence 18, Appl
c 12	276	42.1	466	7	US-10-260-238-4786	Sequence 4786, A
c 13	275.5	42.0	543	7	US-10-021-323-16167	Sequence 16167, A
14	275.5	42.0	779	8	US-10-767-795-271	Sequence 271, App
15	270.5	41.2	507	7	US-10-424-599-5904	Sequence 5904, Ap
16	270.5	41.2	507	7	US-10-425-114-9021	Sequence 9021, Ap
17	266	40.5	636	8	US-10-767-795-5934	Sequence 5934, Ap
18	266	40.5	679	7	US-10-424-599-103390	Sequence 103390,
19	261	39.8	666	7	US-10-424-599-92268	Sequence 92268, A
c 20	259.5	39.6	505	7	US-10-021-323-16005	Sequence 16005, A
21	259.5	39.6	532	7	US-10-021-323-14482	Sequence 14482, A
22	259.5	39.6	538	7	US-10-021-323-11476	Sequence 11476, A
c 23	259.5	39.6	542	7	US-10-021-323-11563	Sequence 11563, A
24	259.5	39.6	604	8	US-10-767-795-5000	Sequence 5000, Ap
25	254.5	38.8	723	8	US-10-425-115-165252	Sequence 165252,
26	249	38.0	577	7	US-10-021-323-13836	Sequence 13836, A
27	247	37.7	737	7	US-10-767-701-20136	Sequence 20136, A
28	246.5	37.6	424	7	US-10-767-701-27843	Sequence 27843, A
29	246	37.5	336	7	US-10-437-963-43786	Sequence 43786, A
30	245	37.3	513	5	US-10-295-403-63	Sequence 63, Appl
31	245	37.3	513	7	US-10-412-6998-661	Sequence 661, App
32	240.5	36.7	613	7	US-10-424-599-50010	Sequence 50010, A
33	239	36.4	273	7	US-10-437-963-48898	Sequence 48898, A
34	239	36.4	776	7	US-10-260-238-433	Sequence 433, App
35	233.5	35.6	980	7	US-10-425-114-26727	Sequence 26727, A
36	233.5	35.6	1056	8	US-10-425-115-39009	Sequence 39009, A
37	228.5	34.8	584	7	US-10-021-323-9298	Sequence 9298, Ap
c 38	228	34.6	932	7	US-10-424-599-56642	Sequence 56642, A
39	227	34.6	1062	8	US-10-425-115-32995	Sequence 32995, A
40	225.5	34.4	717	8	US-10-425-115-5670	Sequence 5670, Ap
41	225.5	34.4	766	8	US-10-425-115-5671	Sequence 5671, Ap
42	222.5	33.9	418	7	US-10-767-701-30624	Sequence 30624, A
43	220.5	33.6	387	7	US-10-767-701-18162	Sequence 18162, A
44	218.5	33.3	330	7	US-10-437-963-55684	Sequence 55684, A
45	214.5	32.7	342	7	US-10-437-963-58875	Sequence 58875, A

ALIGNMENTS

RESULT 1

US-10-697-787-1
; Sequence 1, Application US/10697787
; Publication No. US20040216182A1
; GENERAL INFORMATION:
; APPLICANT: Agrinomics, LLC
; TITLE OF INVENTION: Generation of Plants with Improved Pathogen Resistance and
; FILE OF INVENTION: Brought Tolerance
; FILE REFERENCE: AG03-071C
; CURRENT APPLICATION NUMBER: US/10/697,787
; PRIOR FILING DATE: 2003-10-29
; PRIOR APPLICATION NUMBER: US 60/375,333
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: PCT/US03/12981
; PRIOR FILING DATE: 2003-04-24
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 1
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-697-787-1

Alignment Scores:	
Pred. No.:	1.09e-78
Score:	656.00
Length:	381
Matches:	126
Percent Similarity:	100.0%
Conservative:	0
Best Local Similarity:	100.0%
Mismatches:	0
Query Match:	100.0%
Indels:	0
DB:	8
Gaps:	0

US-10-697-787-2 (1-126) x US-10-697-787-1 (1-381)

QY 1 MetAlaSerAsnSerArgSerSerIleSerProTrrPThrPheSerGlnAsnLysMetPhe 20
Db 1 ATGGCGTCAAACTCAAGAGTTCAATCTCACCATGACGTTTGTAGTCAAAACAAGATGTTTC 60
QY 21 GluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrrPHisAsnValAlaLys 40
Db 61 GAGAGGCGCTTGGCAGTTTACGACAAGGACACACCCGACCGATGGCAATGTCGCAAAA 120
QY 41 AlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAsp 60
Db 121 GCTGTGGAGGAAAACTGTAGAAGAGAGTGAAGCGCCACTATGACATTCTCGTCGAGGAT 180
QY 61 LeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGluSerAsn 80
Db 181 CTCATCAACATCGAGCTGTGTGTCCTTTGCCCAATTACAAGACCTTCGAATCTAAC 240
QY 81 SerArgSerIleAsnAspPheAspThrArgTyrIleThrLysTyrLeuTyrMetMetLeu 100
Db 241 TCAAGAAGCATCAATGACTTTGCACAAAGGTATATAACTAAATATCTATATATGATGCTC 300
QY 101 SerIleTyrPheAspAsnHisSerSerAspPheGluLysPheSerGlnLysValLeuVal 120
Db 301 TCGATATATTTTGATAATCATCTAGTGATTTTGAGAAATTTCTCTCAAAAAGTTCTTGTA 360
QY 121 SerTyrIleSerLeuVal 126
Db 361 AGTTATATTTCTTTGGTT 378

RESULT 2

US-10-512-600-1

; Sequence 1, Application US/10512600

; Publication No. US20050210546A1

; GENERAL INFORMATION:

; APPLICANT: AGRINOMICS, LLC

; TITLE OF INVENTION: GENERATION OF PLANTS WITH IMPROVED PATHOGEN RESISTANCE

; FILE REFERENCE: AG03-033C-US

; CURRENT APPLICATION NUMBER: US/10/512,600

; CURRENT FILING DATE: 2004-10-22

; PRIOR APPLICATION NUMBER: US 60/375,333

; PRIOR FILING DATE: 2002-04-24

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1

; LENGTH: 381

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-10-512-600-1

Alignment Scores:

Pred. No.: 1,09e-78 Length: 381
Score: 656.00 Matches: 126
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0

US-10-697-787-2 (1-126) x US-10-512-600-1 (1-381)

QY 1 MetAlaSerAsnSerArgSerSerIleSerProTrrPThrPheSerGlnAsnLysMetPhe 20
Db 1 ATGGCGTCAAACTCAAGAGTTCAATCTCACCATGACGTTTGTAGTCAAAACAAGATGTTTC 60
QY 21 GluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrrPHisAsnValAlaLys 40
Db 61 GAGAGGCGCTTGGCAGTTTACGACAAGGACACACCCGACCGATGGCAATGTCGCAAAA 120
QY 41 AlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAsp 60
Db 121 GCTGTGGAGGAAAACTGTAGAAGAGAGTGAAGCGCCACTATGACATTCTCGTCGAGGAT 180
QY 61 LeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGluSerAsn 80

Db 181 CTCATCAACATCGAGACTGGTGTGTCCTTTTGCCCAATTACAAGACCTTCGAATCTAAC 240
QY 81 SerArgSerIleAsnAspPheAspThrArgTyrIleThrLysTyrLeuTyrMetMetLeu 100
Db 241 TCAAGAAGCATCAATGACTTTGCACAAAGGTATATAACTAAATATCTATATATGATGCTC 300
QY 101 SerIleTyrPheAspAsnHisSerSerAspPheGluLysPheSerGlnLysValLeuVal 120
Db 301 TCGATATATTTTGATAATCATCTAGTGATTTTGAGAAATTTCTCTCAAAAAGTTCTTGTA 360
QY 121 SerTyrIleSerLeuVal 126
Db 361 AGTTATATTTCTTTGGTT 378

RESULT 3

US-10-424-599-6644

; Sequence 6644, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 6644

; LENGTH: 678

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_106008C.1

US-10-424-599-6644

Alignment Scores:

Pred. No.: 3,76e-29 Length: 678
Score: 293.00 Matches: 69
Percent Similarity: 63.6% Conservative: 20
Best Local Similarity: 49.3% Mismatches: 33
Query Match: 44.7% Indels: 18
DB: 7 Gaps: 4

US-10-697-787-2 (1-126) x US-10-424-599-6644 (1-678)

QY 1 MetAlaSerAsnSer-----ArgSerSerIleSerProTrrPThrPheSerGln 16
Db 206 ATGGCGTCAAGCTCTCTCAGCAAAACAAAGGCTTCTGACTCCTCTTGGAGCCCAAAACAG 265
QY 17 AsnLysMetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrrPHis 36
Db 266 AACAAAGCTGTTGAAAAGCACCTTGCAAAATATGACAAAGGATACCCCTGAGCGTGCAG 325
QY 37 AsnValAlaLysAlaValGlyLysThrValGluGluValLysArgHisTyrAspIle 56
Db 326 AATGTAGCCAAAGCAGTAGGTGCAAAATCTGCAGATGAAGTTAAGACACACTATGAAATC 385
QY 57 LeuValGluAspLeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThr 76
Db 386 CTCTTGGAGGATCTCAGACACATAGAGTCTGGCGCGTTCCTCTTCCCAAGTACAG--- 442
QY 77 PheGluSerAsnSerArgSerIleAsnAspPheAspThrArgTyrIleThrLysTyrLeu 96
Db 443 -----TCCACAGGAAGCAGCACCAATGTTGATGAAGACAGAGGCTTCTGAAGTACTT 496
QY 97 -----TyrMetMetLeuSerIleTyrPheAspAsn 106
Db 497 AAACCTGAATTGATGAAGCAACGTTTTATGTTATCTTCGAGCTATTGGCAGCATTTTATGTGG 556
QY 107 HisSerSerAspPheGluLysPheSerGlnLysValLeu---ValSerTyrIleSerLeu 125

QY 41 AlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluApp 60
Db 267 GCTGTTGGTGGCAAACTCCAGAGGAAGTGAAAGGCACTACGAACCTCTGTGTCAGGAT 326
QY 61 LeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGluSerAsn 80
Db 327 GTTAAGCATATTGAGTCTGGACGTGTGCCATTCCCAAAATTACAAGAAAACTACTTCAGGG 386
QY 81 Ser 81
Db 387 TCA 389

RESULT 7

US-09-770-152-473/c

; Sequence 473, Application US/09770152

; Publication No. US20020040489A1

; GENERAL INFORMATION:

; APPLICANT: Gorlach, Jorn

; APPLICANT: An, Yong-Qiang

; APPLICANT: Hamilton, Carol M.

; APPLICANT: Price, Jennifer L.

; APPLICANT: Raines, Tracy M.

; APPLICANT: Yu, Yang

; APPLICANT: Rameaka, Joshua G.

; APPLICANT: Page, Amy

; APPLICANT: Matthew, Abraham V.

; APPLICANT: Ledford, Brooke L.

; APPLICANT: Woessner, Jeffrey P.

; APPLICANT: Haas, William David

; APPLICANT: Garcia, Carlos A.

; APPLICANT: Kricker, Maja

; APPLICANT: Slader, Ted

; APPLICANT: Davis, Keith R.

; APPLICANT: Allen, Keith

; APPLICANT: Hoffman, Neil

; APPLICANT: Hurban, Patrick

; TITLE OF INVENTION: Expressed Sequences of Arabidopsis

; FILE REFERENCE: thaliana

; CURRENT APPLICATION NUMBER: US/09/770,152

; CURRENT FILING DATE: 2001-01-26

; PRIOR FILING DATE: 2000-01-27

; NUMBER OF SEQ ID NOS: 999

; SOFTWARE: Fast-Seq for Windows Version 4.0

; SEQ ID NO 473

; LENGTH: 556

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-770-152-473

Alignment Scores:

Pred. No.: 2,46e-28 Length: 556

Score: 286.00 Matches: 56

Percent Similarity: 74.7% Conservative: 6

Best Local Similarity: 67.5% Mismatches: 19

Query Match: 43.6% Indels: 2

DB: 3 Gaps: 1

US-10-697-787-2 (1-126) x US-09-770-152-473 (1-556)

QY 1 MetAlaSerAsnSerArgSerSeriIle-----SerProTrrPThrPheSerGlnAsnLys 18
Db 474 ATGCATCAGGCTCAATGCTCTTATGGCTCTGGCTCATGGACTGTTAAGCAGACAAA 415
QY 19 MetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrrPheHisAsnVal 38
Db 414 GCCTTTGAGCGTGTCTAGCAGTCTATGACCAAGACACTCCGGACCGTGTGGCAAAATGTT 355
QY 39 AlalysAlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuVal 58
Db 354 GCTAGAGCTGTGTGGTGAACCAAGAGCTAAGAGACAGTATGACCTTCTAGTT 295

QY 59 GluAspLeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGlu 78
Db 294 CGTGATCGAAGCATCGAGAAATGGTCACGTGCCATTCCCTGACTACAAGACTACTACA 235
QY 79 SerAsnSer 81
Db 234 GGAAACAGC 226

RESULT 8

US-10-225-066A-359

; Sequence 359, Application US/10225066A

; Publication No. US20030226173A1

; GENERAL INFORMATION:

; APPLICANT: Mendel Biotechnology, Inc.

; APPLICANT: RATCLIFFE, Oliver

; APPLICANT: RIECHMANN, Jose Luis

; APPLICANT: ADAM, Luc J

; APPLICANT: DUBELL, Arnold T

; APPLICANT: HEARD, Jacqueline E

; APPLICANT: PILGRIM, Marsha L

; APPLICANT: JIANG, Cai-Zhong

; APPLICANT: REUBER, T. Lynne

; APPLICANT: CREELMAN, Robert A

; APPLICANT: PINEDA, Omdair

; APPLICANT: YU, Guo-Liang

; APPLICANT: BROUN, Pierre E

; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants

; FILE REFERENCE: MBI0036-2 US

; CURRENT APPLICATION NUMBER: US/10/225,066A

; CURRENT FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: 09/837,444

; PRIOR FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: 60/310,847

; PRIOR FILING DATE: 2001-08-09

; PRIOR APPLICATION NUMBER: 60/336,049

; PRIOR FILING DATE: 2001-12-05

; PRIOR APPLICATION NUMBER: 60/338,692

; PRIOR FILING DATE: 2001-12-11

; PRIOR APPLICATION NUMBER: 10/171,468

; PRIOR FILING DATE: 2002-06-14

; NUMBER OF SEQ ID NOS: 1122

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 359

; LENGTH: 578

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-10-225-066A-359

Alignment Scores:

Pred. No.: 2,61e-28 Length: 578

Score: 286.00 Matches: 56

Percent Similarity: 74.7% Conservative: 6

Best Local Similarity: 67.5% Mismatches: 19

Query Match: 43.6% Indels: 2

DB: 6 Gaps: 1

US-10-697-787-2 (1-126) x US-10-225-066A-359 (1-578)

QY 1 MetAlaSerAsnSerArgSerSeriIle-----SerProTrrPThrPheSerGlnAsnLys 18
Db 108 ATGCATCAGGCTCAATGCTCTTATGGCTCTGGCTCATGGACTGTTAAGCAGACAAA 167
QY 19 MetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrrPheHisAsnVal 38
Db 168 GCCTTTGAGCGTGTCTAGCAGTCTATGACCAAGACACTCCGGACCGTGTGGCAAAATGTT 227
QY 39 AlalysAlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuVal 58
Db 228 GCTAGAGCTGTGTGGTGAACCAAGAGCTAAGAGACAGTATGACCTTCTAGTT 287
QY 59 GluAspLeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGlu 78

Db 288 CGTGACATCGAAGACATCGAGAAATGGTCACGTGCCATTCCTGACTACAAAGACTACTACA 347

QY 79 SerAsnSer 81

Db 348 GGAAACAGC 356

RESULT 9

US-10-374-780A-2719

; Sequence 2719, Application US/10374780A

; Publication No. US20040019927A1

; GENERAL INFORMATION:

; APPLICANT: Sherman, Bradley K

; APPLICANT: Riechmann, Jose Luis

; APPLICANT: Jiang, Cai-Zhong

; APPLICANT: Heard, Jacqueline E

; APPLICANT: Haake, Volker

; APPLICANT: Creelman, Robert A

; APPLICANT: Ratcliffe, Oliver

; APPLICANT: Adam, Luc J

; APPLICANT: Reuber, T. Lynne

; APPLICANT: Keddle, James

; APPLICANT: Broun, Pierre E

; APPLICANT: Pilgrim, Marsha L

; APPLICANT: Dubell III, Arnold T

; APPLICANT: Pineda, Omaira

; APPLICANT: Yu, Guo-Liang

; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS

; FILE REFERENCE: MBI-0047 CIP

; CURRENT APPLICATION NUMBER: US/10/374,780A

; CURRENT FILING DATE: 2003-02-25

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

Db 168 GCCTTTGAGCGTCTAGCAGTCTATGACCAAGACACTCCGACCGTTGGCAACAATGTT 227

QY 39 AlaLysAlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuVal 58

Db 228 GCTAGAGCTGTTGGTAAACACACAGAAAGCTTAAGAGACAGTATGACCTTCTAGTT 287

QY 59 GluAspLeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGlu 78

Db 288 CGTGACATCGAAGACATCGAAGTGGTCACGTGCCATTCCTGACTACAAAGACTACTACA 347

QY 79 SerAsnSer 81

Db 348 GGAAACAGC 356

RESULT 10

US-10-225-066A-359

; Sequence 359, Application US/10225066A

; Publication No. US20050160493A9

; GENERAL INFORMATION:

; APPLICANT: Mendel Biotechnology, Inc.

; APPLICANT: Ratcliffe, Oliver

; APPLICANT: Riechmann, Jose Luis

; APPLICANT: Adam, Luc J

; APPLICANT: DUBELL, Arnold T

; APPLICANT: HEARD, Jacqueline E

; APPLICANT: PILGRIM, Marsha L

; APPLICANT: JIANG, Cai-Zhong

; APPLICANT: REUBER, T. Lynne

; APPLICANT: CREELMAN, Robert A

; APPLICANT: PINEDA, Omaira

; APPLICANT: YU, Guo-Liang

; APPLICANT: BROWN, Pierre E

; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants

; FILE REFERENCE: MBI0036-2 US

; CURRENT APPLICATION NUMBER: US/10/225,066A

; CURRENT FILING DATE: 2002-08-09

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2001-04-18

Alignment Scores:
Pred. No.: 2,61e-28 Length: 578
Score: 286.00 Matches: 56
Percent Similarity: 74.7% Conservative: 6
Best Local Similarity: 67.5% Mismatches: 19
Query Match: 43.6% Indels: 2
DB: 9 Gaps: 1

US-10-697-787-2 (1-126) x US-10-225-066A-359 (1-578)

QY 1 MetAlaSerAsnSerArgSerSerile-----SerProTTrpThrPheSerGlnAsnLys 18

Db 108 ATGGCATCAGGCTCAATGCTCTTCTATGGCTCTGGCTCATGACTGTTAAGCAGACAA 167

QY 19 MetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnVal 38

Db 168 GCCTTTGAGCGTCTAGCAGTCTATGACCAAGACACTCCGACCGTTGGCAACAATGTT 227


```
Db 132 CAATTACAAGAAACAGATACATACATATATGTGATG----- 97
QY 108 SerSerAspPheGluLysPheSerGlnLysValLeu---ValSerTyrIleSerLeuVal 126
Db 96 -----AAGTATCAGACATCCCTGCTCTAAGTAGTGTCTGTCTCTGTGTC 52

RESULT 13
US-10-021-323-16167/c
; Sequence 16167, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 16167
; LENGTH: 543
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3829-015-Q1-N6-E2
US-10-021-323-16167

Alignment Scores:
Pred. No.: 6,28e-27 Length: 543
Score: 275.50 Matches: 54
Percent Similarity: 80.0% Conservative: 10
Best Local Similarity: 67.5% Mismatches: 13
Query Match: 42.0% Indels: 3
DB: 7 Gaps: 1

US-10-697-787-2 (1-126) x US-10-021-323-16167 (1-543)

QY 7 SerSerIleSer-----ProTrrPheSerGlnAsnLysMetPheGluArgAla 23
Db 527 TCATCGATTTCATATGCTGTTTCATGCAGCAGCAAGCAAAACAAAGATTTCGAAGGGCT 468
QY 24 LeuAlaValTyrAspLysAspThrProAspArgTTPHisAsnValAlaLysAlaValGly 43
Db 467 TTAGCTGTTTACGACAAGGACACACCATGCTTGTGTACAAATGTTGCTAAAGCTGTGGGA 408
QY 44 GlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAspLeuIleAsn 63
Db 407 GAGAAACTGTTGAGGAGTGAAGAGCACTATGAGCTTCTTTGAAGATGTTAGACGC 348
QY 64 IleGluThrGlyArgValProLeuProLeuProAsnTyrLysThrPheGluSerAsnSerArgSer 83
Db 347 ATCGAGTCGGGTCGGGTTCTTTCCTCCGACTATTGCGGCAATCGTCAAGCA 288

RESULT 14
US-10-767-795-271
; Sequence 271, Application US/10767795
; Publication No. US20040181830A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53534)B
; CURRENT APPLICATION NUMBER: US/10/767,795
; PRIOR FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 117596
; SEQ ID NO 271
; LENGTH: 779
```

```
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C42883_1
US-10-767-795-271

Alignment Scores:
Pred. No.: 1,09e-26 Length: 779
Score: 275.50 Matches: 54
Percent Similarity: 80.0% Conservative: 10
Best Local Similarity: 67.5% Mismatches: 13
Query Match: 42.0% Indels: 3
DB: 8 Gaps: 1

US-10-697-787-2 (1-126) x US-10-767-795-271 (1-779)

QY 7 SerSerIleSer-----ProTrrPheSerGlnAsnLysMetPheGluArgAla 23
Db 150 TCATCGATTTCATATGCTGTTTCATGCAGCAGCAAGCAAAACAAAGATTTCGAAGGGCT 209
QY 24 LeuAlaValTyrAspLysAspThrProAspArgTTPHisAsnValAlaLysAlaValGly 43
Db 210 TTAGCTGTTTACGACAAGGACACACCATGCTTGTGTACAAATGTTGCTAAAGCTGTGGGA 269
QY 44 GlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAspLeuIleAsn 63
Db 270 GAGAAACTGTTGAGGAGTGAAGAGCACTATGAGCTTCTTTGAAGATGTTAGACGC 329
QY 64 IleGluThrGlyArgValProLeuProLeuProAsnTyrLysThrPheGluSerAsnSerArgSer 83
Db 330 ATCGAGTCGGGTCGGGTTCTTTCCTCCGACTATTGCGGCAATCGTCAAGCA 389

RESULT 15
US-10-424-599-5904
; Sequence 5904, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 5904
; LENGTH: 507
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_10533C.1
US-10-424-599-5904

Alignment Scores:
Pred. No.: 2,69e-26 Length: 507
Score: 270.50 Matches: 55
Percent Similarity: 66.7% Conservative: 13
Best Local Similarity: 53.9% Mismatches: 31
Query Match: 41.2% Indels: 3
DB: 7 Gaps: 1

US-10-697-787-2 (1-126) x US-10-424-599-5904 (1-507)

QY 3 SerAsnSerArgSerSerIleSerProTrrPheSerGlnAsnLysMetPheGluArg 22
Db 42 GCTACAACAATGGCTCAAGCCAGGGTTGGACTCCGAGGACAGACAAGAGATTTTCAGAT 101
QY 23 AlaLeuAlaValTyrAspLysAspThrProAspArgTTPHisAsnValAlaLysAlaVal 42
Db 102 GCCCTTGCATCTTTCGACAAGGACACCCACAGATAGTGGCAGACCGTGGCCAGGCCCGTC 161
```

```

Qy      43  GlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAspLeuIle 62
Db      162  GGAGGAAAAACGGTGGAGGAAGTGAAAAAGGCATTATGAGAAAGCTCGTGGAAAGATGTCAAG 221

Qy      63  AsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGluSerAsnSerArg 82
Db      222  AAGATAGAGGAAGGTCAACGTGCCCTCCCCCAATTACCGA-----AGTGCTGCAAGA 272

Qy      83  SerIleAsnAspPheAspThrArgTyrIleThrLysTyrLeuTyrMetMetLeuSerIle 102
Db      273  GGCTACGGTTACATGGATGAAGAACACAGGATGAGGCTCTGAGCTGCGCAGTGAAGTGTG 332

Qy      103  TyrPhe 104
Db      333  ATTTAC 338

```

Search completed: March 10, 2006, 19:49:25
 Job time : 659 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 19:26:25 ; Search time 151 Seconds
(without alignments)
1483.263 Million cell updates/sec

Title: US-10-697-787-2

Perfect score: 656

Sequence: 1 MASNRRSSISPTWFSQNMKFM.....HSSDFEKFQKLVSVISLV 126

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xl

-MODEL=frame+P2n.model -DEV=xl
-Q/abs/ABSSWEB_spool/US10697787/runat_09032006_150631_3419/app_query.fasta_1
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs05p -USER=US10697787 @CGN 1.1.290 @runat_09032006_150631_3419
-NCPU=6 -ICPU=3 -NO MMAP -NEG SCORE=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*

1: /cgn2_6/prodata/1/ina/1 COMB.seq:*
2: /cgn2_6/prodata/1/ina/5 COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B COMB.seq:*
5: /cgn2_6/prodata/1/ina/H COMB.seq:*
6: /cgn2_6/prodata/1/ina/PCTUS COMB.seq:*
7: /cgn2_6/prodata/1/ina/PP COMB.seq:*
8: /cgn2_6/prodata/1/ina/RE COMB.seq:*
9: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	195	29.7	357	3	US-09-640-211A-1581
2	193	29.4	574	3	Sequence 1581, Ap
3	166	25.3	383	3	US-09-640-211A-1534
4	131.5	20.0	532	3	Sequence 1534, Ap
5	124	18.9	1860	3	US-09-640-211A-1310
6	124	18.9	2069	3	Sequence 1284, Ap
7	118	18.0	19112	3	US-09-919-497-50
8	118	18.0	36546	3	Sequence 50, Appl
9	117.5	17.9	345	3	US-09-949-016-5801
					Sequence 5801, Ap
					US-10-697-787-2 (1-126) x US-09-640-211A-1581 (1-357)
					Sequence 17541, A
					Sequence 17543, A
					Sequence 1308, Ap

10	91	13.9	332	3	US-09-640-211A-1824
C 11	85	13.0	494	3	US-09-270-767-31207
C 12	85	13.0	539	3	US-09-270-767-14937
13	84.5	12.9	2274	3	US-09-270-767-10919
14	81	12.3	450	3	US-09-640-211A-2026
15	81	12.3	541	3	US-09-640-211A-255
16	80	12.2	36181	3	US-08-311-731A-120
C 17	79.5	12.1	1830121	3	US-09-557-884-1
C 18	79.5	12.1	1830121	3	US-09-643-930A-1
C 19	79.5	12.1	1830121	3	US-10-150-865-1
20	79	12.0	1612	3	US-09-640-211A-2104
21	77.5	11.8	1971	3	US-09-248-796A-1108
C 22	77	11.7	1086	3	US-09-134-000C-2502
23	76.5	11.7	373	3	US-09-640-211A-1486
24	76.5	11.7	3759	2	US-08-542-921-1
25	76.5	11.7	3759	2	US-08-880-685-1
26	76.5	11.7	3759	2	US-08-880-684-1
27	76	11.6	2502	3	US-09-543-681A-712
28	74	11.3	562	3	US-09-640-211A-1552
29	74	11.3	919	3	US-09-640-211A-485
C 30	73.5	11.2	1664976	3	US-08-916-421B-1
C 31	73.5	11.2	1664976	3	US-09-692-570-1
32	72.5	11.1	921	3	US-09-134-000C-3196
33	72.5	11.1	1815	3	US-09-614-221A-213
C 34	72	11.0	29350	3	US-09-949-016-11963
C 35	72	11.0	29350	3	US-09-949-016-11963
36	71.5	10.9	1710	3	US-09-350-729A-4
37	71.5	10.9	2160	2	US-08-082-849B-30
38	71.5	10.9	2160	6	PCT-US94-01624-30
39	71.5	10.9	2211	3	US-09-350-729A-2
40	71.5	10.9	2292	3	US-09-350-729A-3
41	71.5	10.9	2295	3	US-09-350-729A-1
42	71.5	10.9	2709	2	US-08-021-601-11
43	71.5	10.9	2709	2	US-08-082-849B-11
44	71.5	10.9	2709	6	PCT-US94-01624-11
45	71.5	10.9	4235	2	US-08-021-601-3

ALIGNMENTS

RESULT 1

US-09-640-211A-1581

Sequence 1581, Application US/09640211A

Patent No. 6833446

GENERAL INFORMATION:

APPLICANT: Wood, Marion

APPLICANT: Shenk, Michael A.

APPLICANT: McGrath, Annette

APPLICANT: Glenn, Matthew

TITLE OF INVENTION: Compositions and Methods for the

TITLE OF INVENTION: Modification of Gene Transcription

FILE REFERENCE: 11000.1021C1U

CURRENT APPLICATION NUMBER: US/09/640,211A

CURRENT FILING DATE: 2000-08-16

NUMBER OF SEQ ID NOS: 2368

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1581

LENGTH: 357

TYPE: DNA

ORGANISM: Pinus radiata

US-09-640-211A-1581

ALIGNMENTS

RESULT 1
US-09-640-211A-1581
; Sequence 1581, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1581
; LENGTH: 357
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-1581

Alignment Scores:	2,548-17	Length:	357
Pred. No.:	195.00	Matches:	43
Score:	67.1%	Conservative:	10
Best Local Similarity:	54.4%	Mismatches:	22
Query Match:	29.7%	Indels:	5
DB:	3	Gaps:	1

US-10-697-787-2 (1-126) x US-09-640-211A-1581 (1-357)

Qy 3 SerAenSerArgSerIleSerProtrpThrPheSerGlnAsnLysMetPheGluArg 22

```
; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1310
; LENGTH: 383
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-1310

Alignment Scores:
Pred. No.: 2,4e-13 Length: 383
Score: 166.00 Matches: 29
Percent Similarity: 74.5% Conservative: 9
Best Local Similarity: 56.9% Mismatches: 13
Query Match: 25.3% Indels: 0
DB: Gaps: 0

US-10-697-787-2 (1-126) x US-09-640-211A-1310 (1-383)
Qy 4 AsnSerArgSerSerIleSerProTrrPheSerGlnAsnLysMetPheGluArgAla 23
Db 229 AACGAGGAGGAGGACCGGAAATGGACCCCTGCCGAGAACAAATGTTGGAAGCG 288
Qy 24 LeuAlaValTyrAspLysAspThrProAspArgTrrPheAsnValAlaLysAlaValGly 43
Db 289 CTGGCGGTGCACGATCAGGACACCGCGGATCGGTGGAGTAGGTCGCGCTCGATGATCCCT 348
Qy 44 GlyLysThrValGluGluValLysArgHisTyr 54
Db 349 GGGAGACGGTGGAGGATGTGTTAAGCACTAT 381

RESULT 4
US-09-640-211A-1284
; Sequence 1284, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1284
; LENGTH: 532
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-1284

Alignment Scores:
Pred. No.: 1.82e-08 Length: 532
Score: 131.50 Matches: 28
Percent Similarity: 57.0% Conservative: 17
Best Local Similarity: 35.4% Mismatches: 33
Query Match: 20.0% Indels: 1
DB: Gaps: 3

US-10-697-787-2 (1-126) x US-09-640-211A-1284 (1-532)
Qy 12 TrpThrPheSerGlnAsnLysMetPheGluArgAlaLeuAlaValTyrAspLysAspThr 31
Db 171 TGGAGCCGCTACGAGGACAAAGTCTTCGACACCGCTGTGGCGGTGGCGGAGACTCG 230
Qy 32 ProAspArgTrrPheAsnValAlaLysAlaValGlyLysThrValGluGluValLys 51
Db 231 CCCGACCGGTGGCAGCTGATCGGAAACCGGCTG---AACCGGTCCGCGTCCGAAGTTC 287
```

```
; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1534
; LENGTH: 574
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-1534

Alignment Scores:
Pred. No.: 9.31e-17 Length: 574
Score: 193.00 Matches: 39
Percent Similarity: 66.7% Conservative: 13
Best Local Similarity: 50.0% Mismatches: 26
Query Match: 29.4% Indels: 0
DB: Gaps: 3

US-10-697-787-2 (1-126) x US-09-640-211A-1534 (1-574)
Qy 8 SerIleSerProTrrPheSerGlnAsnLysMetPheGluArgAlaLeuAlaValTyr 27
Db 437 AGTTGTTCTCGTGGAGTAAGGAACAGGACGAGCGGTTCGAGATGCCCTGGCTACCCAC 378
Qy 28 AspLysAspThrProAspArgTrrPheAsnValAlaLysAlaValGlyLysThrVal 47
Db 377 TTCGAGGTTCCCGGATGTTGGAGAGGTAGCGTCCGACGTCGCCGCAAAACCCCTA 318
Qy 48 GluGluValLysArgHisTyrAspIleLeuValGluAspLeuLeuGluThrGly 67
Db 317 GAGGAGATTAAAGCTCCACTACGAGGACCTGGTGGAGATGTCAATAGATAGAGCGTGGC 258
Qy 68 ArgValProLeuProAsnTyrLysThrPheGluSerAsnSerArgSerIleAsn 85
Db 257 CGCGTGCTCTGCGGCTCTACAGTTCGGCTTGGCTCGGATGCTCCGCCAGTAAC 204

RESULT 3
US-09-640-211A-1310
; Sequence 1310, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
```

QY 52 ArgHisTyrAspIleLeuValGluAspLeuIleAsnIleGluThrGlyArgValProLeu 71
Db 288 GAGCACTACCAGAGGCTGGTGGAGGACATTGACCGATCGATCGGGGGGCTCGAGCCG 347
QY 72 ProAsnTyrLysThrPheGluSerAsnSerArgSerIleAsnAspPheAspThrArg 90
Db 348 CCGAGTACCAGCAGCAGCACCACCGCGAGCTGCGGCAGATTGCGCTTCGAGACGAAG 404

RESULT 5

US-09-919-497-50
; Sequence 50, Application US/09919497
; Patent No. 6773883
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50
; LENGTH: 1860
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-497-50

Alignment Scores:

Pred. No.:	1.12e-06	Length:	1860
Score:	124.00	Matches:	21
Percent Similarity:	64.2%	Conservative:	13
Best Local Similarity:	39.6%	Mismatches:	19
Query Match:	18.9%	Indels:	0
DB:	3	Gaps:	0

US-10-697-787-2 (1-126) x US-09-919-497-50 (1-1860)

QY 7 SerSerIleSerProTyrThrPheSerGlnAsnLysMetPheGluArgAlaLeuAlaVal 26
Db 1530 ACAGACTTCACCCCTTGGACACAGAAAGACAGAGCTTTTGGAAACAGCTTTGAAACA 1589
QY 27 TyrAspLysAspThrProAspArgTyrHisAsnValAlaLysAlaValGlyLysThr 46
Db 1590 TACCCAGTAAATACACCTGAAGATGGGAAAAAATAGCAGAAAGCGTGCCTGGCAGGACA 1649
QY 47 ValGluGluValLysArgHisTyrAspIleLeuValGlu 59
Db 1650 AAGAGGAGTGCATGAACAGATACAAAGGAACCTTGTGCAG 1688

RESULT 6

US-09-949-016-5801
; Sequence 5801, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5801
; LENGTH: 2069
; TYPE: DNA

; ORGANISM: Human
US-09-949-016-5801

Alignment Scores:

Pred. No.:	1.3e-06	Length:	2069
Score:	124.00	Matches:	21
Percent Similarity:	64.2%	Conservative:	13
Best Local Similarity:	39.6%	Mismatches:	19
Query Match:	18.9%	Indels:	0
DB:	3	Gaps:	0

US-10-697-787-2 (1-126) x US-09-949-016-5801 (1-2069)

QY 7 SerSerIleSerProTyrThrPheSerGlnAsnLysMetPheGluArgAlaLeuAlaVal 26
Db 1588 ACAGACTTCACCCCTTGGACACAGAAAGACAGAGCTTTTGGAAACAGCTTTGAAACA 1627
QY 27 TyrAspLysAspThrProAspArgTyrHisAsnValAlaLysAlaValGlyLysThr 46
Db 1628 TACCCAGTAAATACACCTGAAGATGGGAAAAAATAGCAGAAAGCGTGCCTGGCAGGACA 1687
QY 47 ValGluGluValLysArgHisTyrAspIleLeuValGlu 59
Db 1688 AAGAGGAGTGCATGAACAGATACAAAGGAACCTTGTGCAG 1726

RESULT 7

US-09-949-016-17541/c
; Sequence 17541, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17541
; LENGTH: 19112
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17541

Alignment Scores:

Pred. No.:	0.0002	Length:	19112
Score:	118.00	Matches:	35
Percent Similarity:	41.7%	Conservative:	28
Best Local Similarity:	23.2%	Mismatches:	59
Query Match:	18.0%	Indels:	30
DB:	3	Gaps:	2

US-10-697-787-2 (1-126) x US-09-949-016-17541 (1-19112)

QY 3 SerAsnSerArgSerSerIleSerProTyrThrPheSerGlnAsnLysMetPheGluArg 22
Db 17658 TCAGGTCCATATACAGACTTCACCCCTTGGACACAGAAAGACAGAGCTTTTGGACAA 17599
QY 23 AlaLeuAlaValTyrAspLysAspThrProAspArgTyrHisAsnValAlaLysAlaVal 42
Db 17598 GCTTTGAAAAACATACCCAGTAAATACACCTGAAGATGGGAAAAAATAGCAGAAAGCGGTG 17539
QY 43 GlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAspLeu 62
Db 17538 CTGGCAGGACAAAGAGGAGTGCATGAACAGATTCAGAGCTTAAGCGCTTAT 17479
QY 63 AsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPhe----- 77

```
Db 17478 AGATTGTTGTAATTTAAATTTAGCTTGGGTGATATAGCTCATATTTTATGTCTATTAAG 17419
Qy 78 -----|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 17418 TGGTTAATAATAGTGTATTTTCAAAACCACTAGTATGGGTCAATATTATGAGAA 17359
Qy 79 SerAsnSerArgSerIleAsnAspPheAspThrArgTyrIleThrLysTyrLeuTyrMet 98
Db 17358 GAAATGAGCAACCTTAGAAACTTTCTAGGCCCTGG-CAGATAATGAAATCAGTATATTTA 17300
Qy 99 MetLeuSer-----|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 17299 GTACTATCTATCTAGTGTGCTGAATCTTTACACCAACACTAGATTTAAATGTCTGTT 17240
Qy 114 PheSerGlnLysValLeuValSerTyrIleSer 124
Db 17239 TTTAAATCGCAACTTCTTAAATCACACCTTTCT 17207
RESULT 8
US-09-949-016-17543
; Sequence 17543, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17543
; LENGTH: 36546
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17543
Alignment Scores:
Pred. No.: 0.000503 Length: 36546
Score: 118.00 Matches: 35
Percent Similarity: 41.7% Conservatives: 28
Best Local Similarity: 23.2% Mismatches: 59
Query Match: 18.0% Indels: 30
DB: 3 Gaps: 2
US-10-697-787-2 (1-126) x US-09-949-016-17543 (1-36546)
Qy 3 SerAsnSerArgSerIleSerProTyrThrPheSerGlnAsnLysMetPheGluArg 22
Db 33729 TCAGGTCATATACAGACTTCACCCCTTGACACAGAAAGACAGAAAGCTTTTGGACAA 33788
Qy 23 AlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnValAlaLysAlaVal 42
Db 33789 GCITTTGAAACATACCCAGTAATAACACCTGAAAGATGGGAAAAAATAGCAGAAAGCGTG 33848
Qy 43 GlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAspLeuLe 62
Db 33849 CCTGGCAGGACAAAGAGGACTGCTAATCAACGATACAAAGTTGCAGACTTAAGCGTTAT 33908
Qy 63 AsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPhe-----77
Db 33909 AGATTGTTGTAATTTAAATTTAGCTTGGGTGATATAGCTCATATTTTATGTCTATTAAG 33968
Qy 78 -----|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 33969 TGGTTAATAATAGTCTTATTTCAAAACCACTAGTATGGGTCAATATTATGAGAA 34028
```

```
Qy 79 SerAsnSerArgSerIleAsnAspPheAspThrArgTyrIleThrLysTyrLeuTyrMet 98
Db 34029 GAAATGAGCAACCTTAGAAACTTTCTAGGCCCTGG-CAGATAATGAAATCAGTATATTTA 34087
Qy 99 MetLeuSer-----|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 34088 GTACTATCTATCTAGTGTGCTGAATCTTTACACCAACACTAGATTTAAATGTCTGTT 34147
Qy 114 PheSerGlnLysValLeuValSerTyrIleSer 124
Db 34148 TTTAAATCGCAACTTCTTAAATCACACCTTTCT 34180
RESULT 9
US-09-640-211A-1308
; Sequence 1308, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; MODIFICATION OF GENE TRANSCRIPTION
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1308
; LENGTH: 345
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-1308
Alignment Scores:
Pred. No.: 7.77e-07 Length: 345
Score: 117.50 Matches: 23
Percent Similarity: 62.1% Conservatives: 13
Best Local Similarity: 39.7% Mismatches: 21
Query Match: 17.9% Indels: 1
DB: 3 Gaps: 1
US-10-697-787-2 (1-126) x US-09-640-211A-1308 (1-345)
Qy 12 TrpThrPheSerGlnAsnLysMetPheGluArgAlaLeuAlaValTyrAspLysAspThr 31
Db 171 TGGAGCCGCTACGAGCAAGGTCTTCGAGCAGCGCGCTGTCGCGGAGGACTCG 230
Qy 32 ProAspArgTrpHisAsnValAlaLysAlaValGlyGlyLysThrValGluGluValLys 51
Db 231 CCCGACCGGTGGCAGCTGATCGGGAACCCGCTG---AACCGTCCCGTCGCAAGTGTTC 287
Qy 52 ArgHisTyrAspIleLeuValGluAspLeuIleAsnIleGluThrGlyArgVal 69
Db 288 GAGCACTACCAAGGCTGTGGAGGACATTGACCGCATCGAGTCGGGGCGGGTC 341
RESULT 10
US-09-640-211A-1824
; Sequence 1824, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; MODIFICATION OF GENE TRANSCRIPTION
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSEQ for Windows Version 4.0
```

```
; SEQ ID NO 1824
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-1824

Alignment Scores:
Pred. No.: 0.00288 Length: 332
Score: 91.00 Matches: 17
Percent Similarity: 84.6% Conservatives: 5
Best Local Similarity: 65.4% Mismatches: 4
Query Match: 13.9% Indels: 0
DB: 3 Gaps: 0

US-10-697-787-2 (1-126) x US-09-640-211A-1824 (1-332)
QY 49 GluValLysArgHisTyrAspIleLeuValGluAspLeuIleAsnIleGluThrGlyArg 68
Db 4 GAGTGGAGGAGCATACGAGCTCTTGTGTGAGGATGTGACTGTGATTGATGCTGCCCG 63
QY 69 ValProLeuProAsnTyr 74
Db 64 GTTGCTTTCCTGCTTAT 81

RESULT 11
US-09-270-767-31207/c
; Sequence 31207, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 1999-03-17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31207
; LENGTH: 494
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-31207

Alignment Scores:
Pred. No.: 0.033 Length: 494
Score: 85.00 Matches: 23
Percent Similarity: 45.8% Conservatives: 6
Best Local Similarity: 37.1% Mismatches: 29
Query Match: 13.0% Indels: 4
DB: 3 Gaps: 2

US-10-697-787-2 (1-126) x US-09-270-767-31207 (1-494)
QY 2 AlaSerAsnSerArgSerSerIleSerPro-----TrpThrPheSerGlnAsnLys 18
Db 191 GCTTCAGCGGAGGAGCATGTCTAATTCGGAGACCACTGGAGCGCAAGAGCAGCAGCGC 132
QY 19 MetPheGluArgAlaLeuAlaValTyrAspLysAspThr---ProAspArgTrpHisAsn 37
Db 131 GCCCTGGAGCGCGCATAGTCAAGTACCGGAGAGACGGCGCGGAGACCGTTGGCAGAG 72
QY 38 ValAlaLysAlaValGlyLysThrValGluGluValLysArgHisTyrAspIleLeu 57
Db 71 ATCGCCAAACAGTGTCCGAGAGAACCAAGAGGAGTGCCTGGTGGCTACCAAGTATCTC 12
QY 58 ValGlu 59
Db 11 TCGGAG 6

RESULT 12
US-09-270-767-14937/c
; Sequence 14937, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 1999-03-17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14937
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-14937

Alignment Scores:
Pred. No.: 0.0373 Length: 539
Score: 85.00 Matches: 23
Percent Similarity: 46.8% Conservatives: 6
Best Local Similarity: 37.1% Mismatches: 29
Query Match: 13.0% Indels: 4
DB: 3 Gaps: 2

US-10-697-787-2 (1-126) x US-09-270-767-14937 (1-539)
QY 2 AlaSerAsnSerArgSerSerIleSerPro-----TrpThrPheSerGlnAsnLys 18
Db 236 GCTTCAGCGGAGGAGCATGTCTAATTCGGAGACCACTGGAGCGCAAGAGCAGCAGCGC 177
QY 19 MetPheGluArgAlaLeuAlaValTyrAspLysAspThr---ProAspArgTrpHisAsn 37
Db 176 GCCCTGGAGCGCGCATAGTCAAGTACCGGAGAGACGGCGCGGAGACCGTTGGCAGAG 117
QY 38 ValAlaLysAlaValGlyLysThrValGluGluValLysArgHisTyrAspIleLeu 57
Db 116 ATCGCCAAACAGTGTCCGAGAGAACCAAGAGGAGTGCCTGGTGGCTACCAAGTATCTC 57
QY 58 ValGlu 59
Db 56 TCGGAG 51

RESULT 13
US-09-270-767-10919
; Sequence 10919, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 1999-03-17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10919
; LENGTH: 2274
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: n means any nucleotide
US-09-270-767-10919

Alignment Scores:
Pred. No.: 0.338 Length: 2274
Score: 84.50 Matches: 17
Percent Similarity: 53.8% Conservatives: 11
Best Local Similarity: 32.7% Mismatches: 23
Query Match: 12.9% Indels: 1
DB: 3 Gaps: 1

US-10-697-787-2 (1-126) x US-09-270-767-10919 (1-2274)
QY 12 TrpThrPheSerGlnAsnLysMetPheGluArgAlaLeuAlaValTyrAspLysAspThr 31
Db 1924 TGGACCAAGGAGGAGCAAGCCCTGCTCGAAGAGCGCATTAATAAACCTATCAACGACGACA 1983
```

```
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14937
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-14937
```

```
Alignment Scores:
Pred. No.: 0.0373 Length: 539
Score: 85.00 Matches: 23
Percent Similarity: 46.8% Conservatives: 6
Best Local Similarity: 37.1% Mismatches: 29
Query Match: 13.0% Indels: 4
DB: 3 Gaps: 2
```

```
US-10-697-787-2 (1-126) x US-09-270-767-14937 (1-539)
```

```
QY 2 AlaSerAsnSerArgSerSerIleSerPro-----TrpThrPheSerGlnAsnLys 18
Db 236 GCTTCAGCGGAGGAGCATGTCTAATTCGGAGACCACTGGAGCGCAAGAGCAGCAGCGC 177
QY 19 MetPheGluArgAlaLeuAlaValTyrAspLysAspThr---ProAspArgTrpHisAsn 37
Db 176 GCCCTGGAGCGCGCATAGTCAAGTACCGGAGAGACGGCGCGGAGACCGTTGGCAGAG 117
QY 38 ValAlaLysAlaValGlyLysThrValGluGluValLysArgHisTyrAspIleLeu 57
Db 116 ATCGCCAAACAGTGTCCGAGAGAACCAAGAGGAGTGCCTGGTGGCTACCAAGTATCTC 57
QY 58 ValGlu 59
Db 56 TCGGAG 51
```

```
RESULT 13
US-09-270-767-10919
; Sequence 10919, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10919
; LENGTH: 2274
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: n means any nucleotide
US-09-270-767-10919
```

```
Alignment Scores:
Pred. No.: 0.338 Length: 2274
Score: 84.50 Matches: 17
Percent Similarity: 53.8% Conservatives: 11
Best Local Similarity: 32.7% Mismatches: 23
Query Match: 12.9% Indels: 1
DB: 3 Gaps: 1
```

```
US-10-697-787-2 (1-126) x US-09-270-767-10919 (1-2274)
```

```
QY 12 TrpThrPheSerGlnAsnLysMetPheGluArgAlaLeuAlaValTyrAspLysAspThr 31
Db 1924 TGGACCAAGGAGGAGCAAGCCCTGCTCGAAGAGCGCATTAATAAACCTATCAACGACGACA 1983
```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 19:19:36 ; Search time 3742 Seconds
(without alignments)
1575.407 Million cell updates/sec

Title: US-10-697-787-2

Perfect score: 656

Sequence: 1 MASNSRSSISPTWFSQKMF.....HSSDFKFSQKLVLSVLSV 126

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlp
-Q-/abs/ABSSWEB_spool/US10697787/runat_09032006_150629_3371/app_query.fasta_1
-DB=EST -Qfmt=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs06p
-USER=US10697787 @CGN 1.1 8010 @runat_09032006_150629_3371 -NCFU=6 -ICPU=3
-NO WMAP -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gsa1:*
10: gb_gsa2:*
11: gb_gsa3:*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	311	47.4	446	8	L38243 BNAF0581E M
2	308	47.0	728	5	BU890694 P040E03 P
3	305	46.5	626	6	CB920052 VVD058F11
4	305	46.5	662	6	CB919185 VVD042D03
5	304	46.3	474	5	BU668323 MC01026A1
6	299	45.6	349	5	BU822353 UB36DPB08
7	299	45.6	624	7	CV232994 WS0199.B2

C	8	299	45.6	627	7	CV230387	WS01916.B
	9	297	45.3	256	9	CC179487	SALK_0699
	10	294	44.8	598	1	AJ768574	ADJ768574
	11	293.5	44.7	629	8	DR994824	MDas9007G
	12	292	44.5	456	7	CO052301	MdFw20570
C	13	290.5	44.3	463	1	AV440717	AV440717
C	14	290	44.2	787	7	CV281695	WS0182.B2
C	15	289	44.1	906	8	DR559298	WS02616.B
C	16	288.5	44.0	407	1	AV442619	AV442619
	17	288	43.9	497	6	CF603713	BACCA01.0
C	18	288	43.9	611	6	CF373337	CSECS086D
	19	287.5	43.8	513	7	CO997831	pan01-16m
	20	286	43.6	497	1	AI993653	701496923
C	21	285.5	43.5	516	7	CV278823	WS0147.B2
C	22	285.5	43.5	588	7	CV283381	WS0187.B2
C	23	285.5	43.5	662	7	CV277080	WS0142.B2
	24	285	43.4	457	6	CD669972	esp1C.pk0
	25	283.5	43.2	599	8	CV882244	Mdlv3.403
C	26	283.5	43.2	653	8	DR997716	MdFb8004M
	27	283.5	43.2	659	8	DR990507	Mdlr7005H
	28	283.5	43.2	660	8	DR995864	MDas9011J
C	29	283.5	43.2	688	7	CV186968	Mdlv2.401
	30	283	43.1	663	5	BU091680	8t76a07.Y
C	31	282	43.0	476	6	CB035850	VVA008A06
	32	282	43.0	492	6	CB002633	VVB019A08
	33	282	43.0	492	6	CB002772	VVB020F07
	34	282	43.0	492	6	CB004008	VVB034G07
	35	282	43.0	509	6	CB344615	CA48EN000
	36	282	43.0	513	6	CB342316	CA32EN000
	37	282	43.0	528	6	CAB11699	CA41LN021
	38	282	43.0	585	3	BM437293	VVA017C08
	39	282	43.0	618	6	CB344529	CA48EN000
	40	282	43.0	620	6	CB343867	CA48EN000
	41	282	43.0	623	6	CB343483	CA32EN000
	42	282	43.0	624	6	CB341211	CA32EN000
C	43	282	43.0	636	6	CB343539	CA32EN000
	44	282	43.0	637	6	CAB11769	CA41LN031
	45	282	43.0	697	6	CB343846	CA32EN000

ALIGNMENTS

RESULT 1

L38243

L38243 BNAF0581E Mustard flower buds Brassica rapa cdNA, mRNA linear EST 03-JUL-1995

L38243

L38243.1 GI:887283

EST.

Brassica rapa (Brassica campestris)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE

1 (bases 1 to 446)

AUTHORS

Lim,C.O., Kim,H.Y., Kim,M.G., Lee,S.I., Chung,W.S., Park,S.H., Hwang,I. and Cho,M.J.

TITLE

Expressed sequence tags of Chinese cabbage flower bud cDNA

JOURNAL

Plant Physiol. 111 (2), 577-588 (1996)

PUBMED

8787028

COMMENT

Contact: Lim,C.O., Kim,M.G., Hwang,I. and Cho,M.J.

FEATURES

source

Location/Qualifiers
1..446
/organism="Brassica rapa"
/mol_type="mRNA"
/strain="pekinensis"
/db_xref="taxon:3711"
/clone_lib="Mustard flower buds"
/note="Devel_stage = flower bud"

ORIGIN


```

DB:
US-10-697-787-2 (1-126) x CB920052 (1-626)
  1 MetaLaser-----AnSerArgSerSerIleSerProTTPThrPheSerGlnAsn 17
  148 ATGCATCCAGCTCAATGTCCTCCCGCAGCTCTGGCTCATCTCTGGACTGCCAAGCAAAAC 207
  18 LysMetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTTPHHisAsn 37
  208 AAAGCCTTTGAAGAGCGCTTCAGTGTATGACAAAGACACCCCTGACCGCTGTACAAAT 267
  38 ValAlaLysAlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeu 57
  268 GTTGCAGGGCTGTGGTGGGAAACCGTGGAGGAAGTGAAGAGGCACTATGAGATCCTT 327
  58 ValGluAspLeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPhe 77
  328 GTTGAGATATCAAGAGCATCGACTCAGACAAAGTGCCTTCCCCAATTACAAAGACTACC 387
  78 GluSerAsnSerArgSer---IleAsnAspPheAspThrArg 90
  388 GGAGCTAGTGGCGGTCCTCAACATGACGACAGCAAGAGG 429

CB919185      662 bp      mRNA      linear      EST 25-APR-2003
VVD042D03 349003 An expressed sequence tag database for abiotic
stressed berries of Vitis vinifera var. Chardonnay Vitis vinifera
cDNA clone VVD042D03 5, mRNA sequence.
CB919185      1 GI:30133846
EST.
KEYWORDS
SOURCE
ORGANISM
Vitis vinifera
Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; Vitaceae; Vitis.
1 (bases 1 to 662)
Cushman,J.C.
An expressed sequence tag database for abiotic stressed berries of
Vitis vinifera var. Chardonnay
Unpublished (2002)
Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
PCR Primers
FORWARD: T3 20mer
BACKWARD: T7 21mer (backward)
Plate: 042 row: D column: 03
Seq primer: T3 20mer
High quality sequence stop: 662.
FEATURES
source
1..662
/organism="Vitis vinifera"
/mol_type="mRNA"
/db_xref="taxon:29760"
/clone="VVD042D03"
/tissue_type="berries"
/dev_stages="mixed; 8, 9, 11, 13, 15, 16 weeks daf"
/clone_lib="An expressed sequence tag database for abiotic
stressed berries of Vitis vinifera var. Chardonnay".
/note="Vector: Lambda Uni-Zap XR, Bluescript SK; Site_1:
EcoRI; Site_2: XhoI"

ALIGNMENT SCORES
Pred. No.: 2
Score: 2.57e-30      Length: 662
305.00      Matches: 62

ORIGIN
Alignment Scores:
Pred. No.:
Score:

```

```

Percent Similarity: 80.9%      Conservative: 14
Best Local Similarity: 66.0%      Mismatches: 14
Query Match: 46.5%      Indels: 4
DB: Gaps: 2

US-10-697-787-2 (1-126) x CB919185 (1-662)
  1 MetaLaser-----AnSerArgSerSerIleSerProTTPThrPheSerGlnAsn 17
  148 ATGCATCCAGCTCAATGTCCTCCCGCAGCTCTGGCTCATCTCTGGACTGCCAAGCAAAAC 207
  18 LysMetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTTPHHisAsn 37
  208 AAAGCCTTTGAAGAGCGCTTCAGTGTATGACAAAGACACCCCTGACCGCTGTACAAAT 267
  38 ValAlaLysAlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeu 57
  268 GTTGCAGGGCTGTGGTGGGAAACCGTGGAGGAAGTGAAGAGGCACTATGAGATCCTT 327
  58 ValGluAspLeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPhe 77
  328 GTTGAGATATCAAGAGCATCGACTCAGACAAAGTGCCTTCCCCAATTACAAAGACTACC 387
  78 GluSerAsnSerArgSer---IleAsnAspPheAspThrArg 90
  388 GGAGCTAGTGGCGGTCCTCAACATGACGACAGCAAGAGG 429

RESULT 5
BU668323
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Sesamum indicum (sesame)
Sesamum indicum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Lamiales; Pedaliaceae; Sesamum.
1 (bases 1 to 474)
Suh,M.C.
Comparative analysis of Expressed Sequence Tags between Sesamum
indicum and Arabidopsis thaliana developing seeds
Unpublished (2003)
Contact: Suh, Mi Chung
Graduate School of Biotechnology, Korea University
1, 5-Ka, Anam-dong, Sungbuk-ku, Seoul 136-701, Rep. of Korea
Tel: 82 2 3290 3169
Fax: 82 2 927 9028
Email: michung@korea.ac.kr
Plate: 026 row: A column: 10.
FEATURES
source
Location/Qualifiers
1..474
/organism="Sesamum indicum"
/mol_type="mRNA"
/db_xref="taxon:4182"
/tissue_type="developing seed"
/cell_lines="yangbaeck"
/dev_stages="5 to 25 days after flowering"
/clone_lib="MC01"
/note="Vector: pBluescript SK; Site_1: EcoRI; Site_2:
XhoI"

ORIGIN
Alignment Scores:
Pred. No.: 2
Score: 2.24e-30      Length: 474
304.00      Matches: 56
80.2%      Conservative: 13
Best Local Similarity: 65.1%      Mismatches: 13
Query Match: 46.3%      Indels: 4
DB: Gaps: 1

US-10-697-787-2 (1-126) x BU668323 (1-474)

```

```

Qy 1 MetAlaSerAsnSer-----ArgSerSerIleSerProTrpThrPheSerGln 16
Db 162 ATGGCGTCGAGCTCCATGACATCGGACGCTGGCTCAGCCTCAGCGTGACGCTCTAAGCAA 221
Qy 17 AsnLysMetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTTrpHis 36
Db 222 AACAAGCAGTTTCGAGGAGGCTTTGGCAATGTACGACCAAGACACGCGTGAACCTGGCAT 281
Qy 37 AsnValAlaLysAlaValGlyGlyLysThrValGluValLysArgHisTyrAspIle 56
Db 282 AATATACCCGAGCGCTGGTGGAGTACGAGAGGAGTGAAGAGCATTTATGAGATT 341
Qy 57 LeuValGluAspLeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThr 76
Db 342 CTGTTAAGACATTATGACATCGACGCGATCAGGTGCGCAATACCTAATCTACAGGCC 401
Qy 77 PheGluSerAsnSerArg 82
Db 402 ATGGGAGCAACAGCAGA 419

RESULT 6
LOCUS BU822353
DEFINITION BU822353 349 bp mRNA linear EST 15-OCT-2002
5 prime, mRNA sequence.
ACCESSION BU822353
VERSION BU822353.1 GI:23990180
KEYWORDS EST.
SOURCE Populus tremula
ORGANISM Populus tremula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.
1 (bases 1 to 349)
Uneberg,P., Bhalerac,R.R., Jansson,S. and Sterky,F.
The poplar tree transcriptome: Analysis of expressed sequence tags
from multiple libraries
Unpublished (2002)
JOURNAL Umea Plant Science Center
COMMENT Department of Plant Physiology
University of Umea, 901 87 Umea, Sweden
Tel: +46 90 786 5279
Fax: +46 90 786 6676
Email: rupali.bhalerac@plantphys.umu.se.

FEATURES
    source
    1..349
    /organism="Populus tremula"
    /mol_type="mRNA"
    /db_xref="taxon:113636"
    /tissue_type="cambium"
    /clone_lib="Populus tremula cambium cDNA library"

ORIGIN
Alignment Scores:
Pred. No.: 7.03e-30 Length: 349
Score: 299.00 Matches: 56
Percent Similarity: 80.5% Conservative: 10
Best Local Similarity: 68.3% Mismatches: 16
Query Match: 45.8% Indels: 0
DB: 5 Gaps: 0

US-10-697-787-2 (1-126) x BU822353 (1-349)

Qy 3 SerAsnSerArgSerSerIleSerProTrpThrPheSerGlnAsnLysMetPheGluArg 22
Db 99 ACGTCTTCAATGGCTCCGGCTCTCTCTGGACACGCCAAACAAGCTATTTCGAGAAG 158
Qy 23 AlaLeuAlaValTyrAspLysAspThrProAspArgTTrpHisAsnValAlaLysAlaVal 42
Db 159 GCCCTGGCTGTATACGACAAAGACACCCCGACCGCTGGCAAAATGTGGCCAGCCGCTG 218

```

```

Qy 43 GlyGlyLysThrValGluValLysValLysArgHisTyrAspIleLeuValGluAspLeuIle 62
Db 219 GGTGCAAGTCTCTCTGGAAGAGTTAAGAGGCACTATGATCGTCTCTGTGAAGATCTCGTG 278
Qy 63 AsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGluSerAsnSerArg 82
Db 279 TACATAAATCCGGCGAAGCGCTCTGCGAATTACAAAGCCCTCTCGCAGCAATGGTAAA 338
Qy 83 SerIle 84
Db 339 GCAGTT 344

RESULT 7
LOCUS CV232994/c
DEFINITION CV232994 624 bp mRNA linear EST 21-SEP-2004
WS0199 E24 3', mRNA sequence.
ACCESSION CV232994
VERSION CV232994.1 GI:52389099
KEYWORDS EST.
SOURCE Populus trichocarpa (Populus balsamifera subsp. trichocarpa)
ORGANISM Populus trichocarpa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.
1 (bases 1 to 624)
Ralph,S., Cooper,D., Kolosova,N., Oddy,C., Butterfield,Y.,
Kirkpatrick,R., Liu,J., Palmquist,D., Stott,J., Barber,S., Yang,G.,
Babakaiff,R., Brown-John,M., Chand,S., Featherstone,R., Masson,A.,
Mayo,M., Moran,J., Olson,T., Wong,D., Ritland,C.E., Siddiqui,A.,
Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C., Ritland,K. and
Bohlmann,J.
The poplar transcriptome: Analysis of expressed sequence tags from
multiple cDNA libraries
Unpublished (2004)
JOURNAL Contact: Joerg Bohlmann
COMMENT Genome BC forest genomics program
University of British Columbia
Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
Vancouver, British Columbia, Canada, V6T 1Z3
Tel: 1-604-822-0282
Fax: 1-604-822-2114
Email: bohlmann@msl.ubc.ca
Plate: WS0199 row: E column: 24
High quality sequence stop: 624
POLYA=Yes.

FEATURES
    Location/Qualifiers
    1..624
    /organism="Populus trichocarpa"
    /mol_type="mRNA"
    /cultivar="VT-125"
    /db_xref="taxon:3694"
    /clone="WS0199 E24"
    /sex="Not determined"
    /lab_host="E. coli DH10B T1 phage resistant cells"
    /clone_lib="PT-DX-N-A-10"
    /notes="Vector: pBluescript II SK (+) XR; Site_1: EcoRI (5'
    end of cDNA); Site_2: XhoI (3' end of cDNA); Outer xylem
    from 5 year old trees harvested every two weeks between
    April and October of 2002 at the University of British
    Columbia south campus farm in Vancouver, British Columbia.
    mRNA was isolated from each tissue source independently
    and equal quantities of mRNA from each tissue were then
    pooled. cDNA was prepared from 5 micrograms of mRNA and
    directionally ligated into the pBluescript II SK (+) XR
    vector using the pBluescript II XR cDNA Library
    Construction Kit according to manufacturer's instructions
    with modifications (Stratagene). Plasmid DNA was then
    transformed by electroporation into DH10B cells
    (Invitrogen) for propagation. Normalization was applied
    according to published methods [Bonaldo M.F. et al. (1996)
    Genome Research 6(9):791] in order to reduce the abundance
    of highly expressed transcripts."

```

ORIGIN

Alignment Scores:
Pred. No.: 1.53e-29 Length: 624
Score: 299.00 Matches: 56
Percent Similarity: 80.5% Conservative: 10
Best Local Similarity: 68.3% Mismatches: 16
Query Match: 45.6% Indels: 0
DB: 7 Gaps: 0

US-10-697-787-2 (1-126) x CV232994 (1-624)

QY 3 SerAsnSerArgSerSerIleSerProTrrPheSerGlnAsnLysMetPheGluArg 22
Db 566 ACATCTTCAATGGCACCCTCTCTGGACAGCCCAACAAATAAGCTATTCGAGAAG 507
QY 23 AlaLeuAlaValTyrAspLysAspThrProAspArgTrrPheAsnValAlaVal 42
Db 506 GCCTGGCTGTATACACAAAGACACCCAGACCCGCTGGCAAAATGTGCCAAGGCCGTG 447
QY 43 GlyGlyLysThrValGluValLysArgHisTyrAspIleValGluAspLeu 62
Db 446 GTGGCAATCTCTCTGGAAGTTAAGAGCCTATGATCGTCTCTGGAGATCTCGTG 387
QY 63 AsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGluSerAsnSerArg 82
Db 386 TACATAGAATCGGCCAAGCCCTCTGCGGATTACAGCCCTCTGGCAGCAATGGTAGA 327
QY 83 SerIle 84
Db 326 GGACTT 321

RESULT 8

CV230387/7/c
LOCUS WS01916.B21_F23 PT-DX-N-A-10 Populus trichocarpa cDNA clone
DEFINITION WS01916.F23 3', mRNA sequence.
ACCESSION CV230387
VERSION CV230387.1 GI:52383831
KEYWORDS EST.
SOURCE Populus trichocarpa (Populus balsamifera subsp. trichocarpa)
ORGANISM Populus trichocarpa
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Malpighiales; Salicaceae; Salicaceae; Populus.
1 (bases 1 to 627)
Ralph, S., Cooper, D., Kolosova, N., Oddy, C., Butterfield, Y.,
Kirkpatrick, R., Liu, J., Palmquist, D., Stott, J., Barber, S., Yang, G.,
Babakaiff, R., Brown-John, M., Chand, S., Featherstone, R., Masson, A.,
Mayo, M., Moran, J., Olson, T., Wong, D., Ritland, C.E., Siddiqui, A.,
Holt, R., Jones, S., Marra, M., Ellis, B.E., Douglas, C., Ritland, K. and
Bohmann, J.
The poplar transcriptome: Analysis of expressed sequence tags from
multiple cDNA libraries
Unpublished (2004)
Contact: Joerg Bohmann
Genome BC forest genomics program
University of British Columbia
Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
Vancouver, British Columbia, Canada, V6T 1Z3
Tel: 1-604-822-0282
Fax: 1-604-822-2114
Email: bohlmann@ml.ubc.ca
Plate: WS01916 row: F column: 23
High quality sequence stop: 627
POLYA=Yes.

FEATURES

source
1..627
/organism="Populus trichocarpa"
/mol_type="mRNA"
/cultivar="VT-125"
/db_xref="taxon:3694"
/clone="WS01916_F23"
RESULT 9
CC179487
LOCUS CCI179487
DEFINITION SALK_069941.39.90.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_069941.39.90.x, genomic
survey sequence.
ACCESSION CCI179487
VERSION CCI179487.1 GI:30318038
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 256)
Alonso, J.M., Leisse, T.J., Barajas, P., Chen, H., Cheuk, R.,
Gadrinab, C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L.,
Shinn, P., Zimmermann, J. and Ecker, J.R.
A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
Unpublished (2001)

COMMENT

Contact: Joseph R. Ecker
 Salk Institute Genomic Analysis Laboratory (SIGNAL)
 The Salk Institute for Biological Studies
 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
 Tel: 858 453 4100 x1752
 Fax: 858 558 6379
 Email: ecker@salk.edu
 This is single pass sequence recovered from the left border of
 TDNA. This sequence lies within 300 bases of the 5' end of
 At1g75250.

Class: TDNA tagged.

Location/Qualifiers

FEATURES

source

1..256 /organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/ecotype="Col-0"

/db_xref="taxon:3702"

/clone_lib="Arabidopsis thaliana TDNA insertion lines"

/note="PCR was performed on Arabidopsis thaliana lines

each of which contains one or more TDNA insertion

elements. The resultant fragment for each line was

directly sequenced to determine the genomic sequence at

the site of insertion. Details of the protocols used can

be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN

Alignment Scores:
 Pred. No.: 8.65e-30 Length: 256
 Score: 297.00 Matches: 57
 Percent Similarity: 95.4% Conservative: 5
 Best Local Similarity: 87.7% Mismatches: 3
 Query Match: 45.3% Indels: 0
 DB: 9 Gaps: 0
 US-10-697-787-2 (1-126) x CC179487 (1-256)

Qy 1 MetAlaSerAenSerArgSerSerIleSerProTrpThrPheSerGlnAenLysMetPhe 20
 Db 60 ATGGCGTCAAACTCAAAAGTTCAATCTCACATGGAGCGTTAGTCAAAACAAAGATGTC 119
 Qy 21 GluArgAlaLeuAlaValTyAspLysAspThrProAspArgTTPHisAenValAlaLys 40
 Db 120 GAGAGGGCTTGGCAGTTTACGACAGGACACACCCGACCATGGCCAAATGGCCAAA 179
 Qy 41 AlaValGlyLysThrValGluGluValLysArgHisTyAspIleLeuValGluAsp 60
 Db 180 GCTGTCGGAGGAAACATGTATAGAGTGAACCGCTCTATGACATTCGTCCTCAAGAA 239
 Qy 61 LeuIleAenIleGlu 65
 Db 240 CTCATCAACATCCAG 254

RESULT 10
 AJ768574 598 bp mRNA linear EST 06-AUG-2004
 LOCUS AJ768574 Populus euphratica leaf adult Populus euphratica cDNA
 DEFINITION clone P0000100019D06F1, mRNA sequence.

ACCESSION AJ768574
 VERSION AJ768574.1 GI:50061401
 KEYWORDS EST.
 SOURCE Populus euphratica
 ORGANISM Populus euphratica
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids I; Malpighiales; Salicaceae; Salicaceae; Populus.

REFERENCE 1 (bases 1 to 598)
 Brosche, M., Alacalo, E.R., Vinocur, B., Altman, A., Teichmann, T.,
 Ottow, E.A., Polle, A., Djilianov, D., Afif, D., Triboulot, M.B.,
 Dreyer, E., Paulin, L. and Kangasjaervi, J.
 Gene expression in desert leaves
 Unpublished (2004)
 Contact: Prof. Jaakko Kangasjaervi

Department of Biosciences, Plant Physiology
 University of Helsinki
 FIN-00014 Helsinki, Finland.

FEATURES

source

1..598 /organism="Populus euphratica"

/mol_type="mRNA"

/db_xref="taxon:75702"

/clone="P0000100019D06F1"

/tissue_type="leaf"

/dev stage="adult"

/clone_lib="Populus euphratica leaf adult"

/note="country: Israel;Ein Avdat"

ORIGIN

Alignment Scores:
 Pred. No.: 6.86e-29 Length: 598
 Score: 294.00 Matches: 56
 Percent Similarity: 80.0% Conservative: 12
 Best Local Similarity: 65.9% Mismatches: 15
 Query Match: 44.8% Indels: 2
 DB: 1 Gaps: 1
 US-10-697-787-2 (1-126) x AJ768574 (1-598)

Qy 1 MetAlaSerAenSerArgSerSerIleSerProTrpThrPheSerGlnAenLys 18
 Db 160 ATGGCATCAAGCTCAATGCTCTCGCGTTCGGGCTCGTGAGCTCTCAGAGAAACAAA 219
 Qy 19 MetPheGluArgAlaLeuAlaValTyAspLysAspThrProAspArgTTPHisAenVal 38
 Db 220 GCCTTTGAAAGGCTTTGGCCGTGTATGACAGAGACACGCCCGCATCGTGATCAATGTA 279
 Qy 39 AlaLysAlaValGlyLysThrValGluGluValLysArgHisTyAspIleLeuVal 58
 Db 280 GCCAGGCGCTGCTGGAAAAACCGCAGAGAAAGTGAAGAGGCACTATGAATGCTGTGTG 339
 Qy 59 GluAspLeuIleAenIleGluThrGlyArgValProLeuProAenTyLysThrPheGlu 78
 Db 340 GAGGATGTGAAGCATATTGAGTCGGGTGATGTTCTTCCTTAATTACAGGACTACTGGA 399
 Qy 79 SerAenSerArgSer 83
 Db 400 GCAATGGCCACGCA 414

RESULT 11

DR994824

LOCUS

DEFINITION

DR994824 629 bp mRNA linear EST 04-AUG-2005
 Mdaes9007G21.g1 Apple EST Mdaes Malus x domestica cDNA similar to
 ref|NP_179759.1| myb family transcription factor [Arabidopsis
 thaliana] >pir|G84603 hypothetical protein A2921650 [imported] -
 Arabidopsis thaliana >gb|AAD23640.1| unknown protein [Arabidopsis
 thaliana] >gb|AAP40381.1| putative myb family transcrip. mRNA
 sequence.

ACCESSION

DR994824

VERSION

EST.

KEYWORDS

SOURCE

ORGANISM

Malus x domestica

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids I; Rosales; Rosaceae; Maloideae; Malus.

REFERENCE 1 (bases 1 to 629)

Korban, S., Vodkin, L., Liu, L., Gasic, K., Gonzales, O., Hernandez, A.,
 Aldwinckle, H., Malnoy, M., Carroll, N., Goldsborough, P., Orvis, K.,
 Clifton, S., Pape, D., Marra, M., Hillier, L., Martin, J., Wylie, T.,
 Dante, M., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Ronko, I.,
 Tsagareishvili, R., Kennedy, S., Waterston, R. and Wilson, R.

Apple Functional Genomics grant - NSF 0321702

Unpublished (2004)

Contact: Schuyler S. Korban

Apple Functional Genomics grant - NSF 0321702

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

Library material provided by S. Korban/H. Aldwinckle/ M. Malnoy
Washington University Genome Sequencing Center
This trace has been recalled with phred
original value before phred recall for SL was 118
original value before phred recall for SR was 873.

FEATURES

source

1..629

/organism="Malus x domestica"

/mol_type="mRNA"

/cultivar="GoldRush"

/db_xref="taxon:3750"

/tissue_type="Leaves challenged with apple scab (Venturia

inaequalis)"

/lab_host="DH10B ampicillin resistant"

/clone_lib="Apple EST Mdas"

/note="Vector: pBluescript II SK (+); Site 1: XhoI;

Site 2: EcoRI; Total RNA was extracted from freeze dried

leaf tissue, using the method described by Wang and Vodka

(Plant Molecular Biology Reporter 12:132-145, 1994).

Poly(A)+ mRNA was isolated from total RNA using the

PolyAtract mRNA isolation system III (Promega). The

library was prepared using the Stratagene pBluescript II

XR cDNA library construction kit. Complementary DNA was

synthesized from mRNA using a Poly (dT) sequence with a

XhoI restriction site. EcoRI adapters were ligated to the

blunt ended cDNA fragments followed by restriction with

XhoI. The cDNA insert is protected from XhoI digestion via

methylation during first strand cDNA synthesis. The cDNA

fragments were directionally cloned into the EcoRI-XhoI

restriction site of the pBluescript vector. The ligated

cDNA fragments were transformed into E. coli Electromax

DH10B host cells. Transformation efficiency: 1.0E+04

Average insert size by PCR: 1290 bp"

ORIGIN

Alignment Scores:

Pred. No.: 8,578-29 Length: 629
Score: 293.50 Matches: 62
Percent Similarity: 63.8% Conservative: 12
Best Local Similarity: 53.4% Mismatches: 17
Query Match: 44.7% Indels: 25
DB: 8 Gaps: 2

US-10-697-787-2 (1-126) x DR994824 (1-629)

QY 1 MetAlaSer-----AsnSerArgSerSerIleSerProTrrPheSerGlnAen 17
DB 87 ATGGCATCAAGCTCAATGCTTCAAGAGCTCTGGCTCTCTGGACTGCCAAGCAAAC 146
QY 18 LysMetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAen 37
DB 147 AAAGCCTTGAAGAGCTCTGGCTGTCTACGACAAAGGACACTGCTGACCGCTGGTACAT 206
QY 38 ValAlaLysAlaValGlyLysThrValGluGluValLysArgHisTyrAspIleLeu 57
DB 207 GTGGCCAAAGCGGTCCGGCGGCAAAACACCGGAGGAAGTCAAGAGGCATATGAGGTCTT 266
QY 58 ValGluAspLeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPhe 77
DB 267 GTGGAAGATGTCACATATTAGTCAGCCAGGAGGTCCTTCCAGATTACAGGACTACT 326
QY 78 GluSerAsnSerArgSerIleAsnAspPheAspThrArgTyrIleThrLysTyrLeuTyr 97
DB 327 GGTGGGAACAGCCATGGC----- 344
QY 98 MetMetLeuSerIleTyrPheAspAsnHisSerSerAspPheGluLys 113
DB 345 -----CACCATGATGATGAGGAAAG 368

RESULT 12
CO052301

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..456

/organism="Malus x domestica"

/mol_type="mRNA"

/cultivar="GoldRush"

/db_xref="taxon:3750"

/clone="Mdfw2057018"

/tissue_type="Flower"

/lab_host="DH10B ampicillin resistant"

/clone_lib="Mdfw"

/note="Vector: pBluescript II SK (+); Site 1: NotI;

Site 2: EcoRI; Total RNA was extracted separately from

each floral stage (bud, balloon, open and after

pollination), using the 'pine tree' method. Poly(A)+mRNA

was isolated twice from total RNA from each stage using

the Oligotex Direct mRNA kit (Qiagen). mRNA was reverse

transcribed into double stranded cDNA using a modified

oligo18(dT) primer with an identifying tag sequence (see

table below). cDNA's from different stages were pooled in

equal amounts before adaptor ligation. Tag identification

when sequencing from 5' end; Stage 1 (bud) insert

18(A)TCGA; Stage 2 (balloon) insert 18(ATCGCA; Stage 3

(open) insert 18(ATCGGT; Stage 4 (afterpollination)

insert 18(ATCGGT. Tag identification when sequencing from

3' end; Stage 1 (bud) TCGA18(T) insert; Stage 2 (balloon)

TCGCA18(T) insert; Stage 3 (open) ACGCA18(T) insert; Stage

4 (afterpollination) ACCGA18(T) insert. Double stranded

cDNAs were size selected (more than 450 bp), adaptor

with EcoRI adapters at both ends and then digested with

NotI. The cDNAs were then directionally cloned into

EcoRI-NotI digested pBS II SK(+) phagemid

vector(Stratagene). Identification of adaptors and tags in

5'-end sequenced clones: (Vector) TAAGCTT(End

Vector) (Start EcoRI adaptor)GATATCGAATTCATGTTGGG (End

EcoRI adaptor) (Start Tag)TCGA(End Tag) (Start

Insert) (Start Tag)TCGA(End Tag) (Start

NotI site)Vector)GCGCCGCCACCGGG. The total number of

white colony forming units (cfu) in the primary library

CO052301 456 bp mRNA linear EST 14-JUN-2004
Mdfw2057018.y1 Mdfw Malus x domestica cDNA clone Mdfw2057018 5',
similar to TR:Q9SIJ5 Q9SIJ5 AT2G21650 PROTEIN. ; mRNA sequence.

CO052301
CO052301.1 GI:48693766

EST.
Malus x domestica

Malus x domestica

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Rosales; Rosaceae; Maloideae; Malus.

1 (bases 1 to 456)

Korban,S., Vodkin,L., Liu,L., Gasic,K., Gonzales,O., Hernandez,A.,
Aldwinckle,H., Malnoy,M., Carroll,N., Goldsborough,P., Orvis,K.,
Clifton,S., Pape,D., Marra,M., Hillier,L., Martin,J., Wylie,T.,
Dante,M., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Ronko,I.,
Tsagarisvilli,R., Kennedy,S., Waterston,R. and Wilson,R.

Apple Functional Genomics grant - NSF 0321702

Unpublished (2004)

Contact: Schuyler S. Korban

Apple Functional Genomics grant - NSF 0321702

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Library materials provided by: Schuyler S. Korban Library

constructed by: A. Hernandez / K. Gasic Library sequenced by:

Washington University Genome Sequencing Center

WashU EST name: aah2h09.y1

Seq primer: -40UP from Gibco

High quality sequence stop: 456.

Location/Qualifiers

1..456

/organism="Malus x domestica"

/mol_type="mRNA"

/cultivar="GoldRush"

/db_xref="taxon:3750"

/clone="Mdfw2057018"

/tissue_type="Flower"

/lab_host="DH10B ampicillin resistant"

/clone_lib="Mdfw"

/note="Vector: pBluescript II SK (+); Site 1: NotI;

Site 2: EcoRI; Total RNA was extracted separately from

each floral stage (bud, balloon, open and after

pollination), using the 'pine tree' method. Poly(A)+mRNA

was isolated twice from total RNA from each stage using

the Oligotex Direct mRNA kit (Qiagen). mRNA was reverse

transcribed into double stranded cDNA using a modified

oligo18(dT) primer with an identifying tag sequence (see

table below). cDNA's from different stages were pooled in

equal amounts before adaptor ligation. Tag identification

when sequencing from 5' end; Stage 1 (bud) insert

18(A)TCGA; Stage 2 (balloon) insert 18(ATCGCA; Stage 3

(open) insert 18(ATCGGT; Stage 4 (afterpollination)

insert 18(ATCGGT. Tag identification when sequencing from

3' end; Stage 1 (bud) TCGA18(T) insert; Stage 2 (balloon)

TCGCA18(T) insert; Stage 3 (open) ACGCA18(T) insert; Stage

4 (afterpollination) ACCGA18(T) insert. Double stranded

cDNAs were size selected (more than 450 bp), adaptor

with EcoRI adapters at both ends and then digested with

NotI. The cDNAs were then directionally cloned into

EcoRI-NotI digested pBS II SK(+) phagemid

vector(Stratagene). Identification of adaptors and tags in

5'-end sequenced clones: (Vector) TAAGCTT(End

Vector) (Start EcoRI adaptor)GATATCGAATTCATGTTGGG (End

EcoRI adaptor) (Start Tag)TCGA(End Tag) (Start

Insert) (Start Tag)TCGA(End Tag) (Start

NotI site)Vector)GCGCCGCCACCGGG. The total number of

white colony forming units (cfu) in the primary library

before amplification was 1.1x10⁶ cfu (colony forming units). The background of empty clones was less than 1%. Inserts ranged from 0.5kb to 3 kb, as determined by PCR. Purified plasmid DNA from the primary library was converted to single-stranded circles and used as a template for PCR amplification using the T7 and T3 priming sites flanking the cloned cDNA inserts. The purified PCR products, representing the entire cloned cDNA population, were used as a driver for normalization. Hybridization between the single-stranded library and the PCR products was carried out for 44 hours at 30C. Unhybridized single-stranded DNA circles were separated from hybridized DNA remaining partially double-stranded and electroporated into DH10B cells to generate the normalized library. The total number of clones with insert was 9x10⁶ cfu. Background of empty clones was less than 1%."

ORIGIN

Alignment Scores:
Pred. No.: 8,88e-29 Length: 456
Score: 292.00 Matches: 53
Percent Similarity: 87.8% Conservative: 12
Best Local Similarity: 71.6% Mismatches: 9
Query Match: 44.5% Indels: 0
DB: 7 Gaps: 0

US-10-697-787-2 (1-126) x C0052301 (1-456)

QY 3 SerAsnSerArgSerSerIleSerProTyrThrPheSerGlnAsnLysMetPheGluArg 22
DB 60 TCCTCTCGAGGAATTTCACTCTCTCGTGACCGCCGAGGAGAACAGCTGTTTCGAGAG 119
QY 23 AlalaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnValAlaLysAlaVal 42
DB 120 GCGTGGCCGTGTACGACAGGACACGACGACCGCTGGCATAATGTGGCCAGGCGGTC 179
QY 43 GlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAspLeuLe 62
DB 180 GCGGGAATCTCGCGGAGGAAGTAAGAGGACACTACGAGATCTCGTGCATGATCTCATG 239
QY 63 AsnIleGluThrGlyArgValProLeuProAsnTyrLysThr 76
DB 240 CACATAGATCCGGCCATGTCCCTATACCAAAATTAAGTCC 281

RESULT 13
AV440717/c 463 bp mRNA linear EST 18-FEB-2004
LOCUS AV440717 Arabidopsis thaliana above-ground organ two to six-week
DEFINITION Old Arabidopsis thaliana cDNA clone AP207c09_f 3', mRNA sequence.
ACCESSION AV440717
VERSION AV440717.1 GI:7611088
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 463)
AUTHORS Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
JOURNAL DNA Res. 7 (3), 175-180 (2000)
PUBMED 10907847
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
1..463
source /organism="Arabidopsis thaliana"
/mol_type="mRNA"

/scotype="Columbia"
/db_xref="taxon:3702"
/clone="AP207c09_f"
/issue_type="aboveground organs"
/dev_stage="two to six-week old"
/clone_lib="Arabidopsis thaliana above-ground organ two to six-week old"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Alignment Scores:
Pred. No.: 1.45e-28 Length: 463
Score: 290.50 Matches: 59
Percent Similarity: 70.4% Conservative: 22
Best Local Similarity: 51.3% Mismatches: 25
Query Match: 44.3% Indels: 9
DB: 1 Gaps: 2

US-10-697-787-2 (1-126) x AV440717 (1-463)

QY 1 MetAlaSerAsnSerArgSerIleSerProTyrThrPheSerGlnAsnLysMetPhe 20
DB 385 ATGGCTTCCAACCTCAATGAGCTCTAGCGCTTCTTGACACGTAAGGAGAACAAATATTT 326
QY 21 GluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnValAlaLys 40
DB 325 GAAAGGCGGTGGCTACATATGACAGGACACTCTCGACCGTTGGCATACGTTGCAAGA 266
QY 41 AlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAsp 60
DB 265 GCGTGTGGCGCAATACGCTGAAGAAGTAAGCGGACACTACGAGCTCTCATTAGGAT 206
QY 61 LeuIleAlaValGluThrGlyArgValProLeuProAsnTyrLysThrPheGluSerAs 80
DB 205 GTCAATGACATTGAGTCAGGCGGTTATCCACATCCCAATTTACCGTTCAATGAAACAAC 146
QY 80 nSerArgSerIle-AsnAspPheAspThrArgTyrIleThrLysTyrLeuTyrMetMetL 100
DB 145 CACTGAAGCATTTAAAGGAACCTCAA---AAGGCTCTACTGAAGTGG----- 102
QY 100 euSerIleTyrPheAspAsnHisSerSerAspPheGluLys 113
DB 101 -----ACATGGTTCCAAACCATTTCACAAAGCACATCTAAG 66

RESULT 14

CV281695 787 bp mRNA linear EST 22-SEP-2004
LOCUS WS0182_B21_J21 PTXD-II-N-A-9 Populus trichocarpa x Populus
DEFINITION deltoides cDNA clone WS0182_J21 3', mRNA sequence.
ACCESSION CV281695
VERSION CV281695.1 GI:52534670
KEYWORDS EST.
SOURCE Populus trichocarpa x Populus deltoides
ORGANISM Populus trichocarpa x Populus deltoides

REFERENCE 1 (bases 1 to 787)
AUTHORS Ralph, S., Cooper, D., Kolosova, N., Oddy, C., Butterfield, Y., Kirkpatrick, R., Liu, J., Palmquist, D., Stott, J., Barber, S., Yang, G., Babakaiff, R., Brown-John, M., Chand, S., Featherstone, R., Mason, A., Mayo, M., Moran, J., Olson, T., Wong, D., Ritland, C.E., Siddiqui, A., Holt, R., Jones, S., Marra, M., Ellis, B.E., Douglas, C., Ritland, K., and Bohlmann, J.
TITLE The poplar transcriptome: Analysis of expressed sequence tags from multiple cDNA libraries
JOURNAL Unpublished (2004)
COMMENT Contact: Joerg Bohlmann
Genome BC forest genomics program
University of British Columbia
Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
Vancouver, British Columbia, Canada, V6T 1Z3

Tel: 1-604-822-0282
Fax: 1-604-822-2114
Email: bohlmann@sl.ubc.ca
Plate: WS0182 row: J column: 21
High quality sequence stop: 787
POLYA=Yes.

FEATURES

source

1. 787
/organism="Populus trichocarpa x Populus deltoides"
/mol_type="mRNA"
/cultivar="H11-11"
/db_xref="taxon:3695"
/clone="WS0182_J21"
/sex="Male"
/lab_host="E. coli DH10B T1 phage resistant cells"
/clone_lib="PTXD-IL-N-A-9"
/notes="Vector: pBluescript II SK (+) XR; Site 1: EcoRI (5' end of cDNA); Site 2: XhoI (3' end of cDNA); Sapling trees two metres in height and grown under greenhouse conditions were exposed to continuous feeding by Malacosoma disstria Hubner (forest tent caterpillar) mid-instar larvae caged on the sapling using mesh bags. Mature leaves from within the caged region were collected 2 hours, 12 hours, 24 hours and 48 hours after the onset of treatment. mRNA was isolated from each tissue source independently and equal quantities of mRNA from each tissue were then pooled. cDNA was prepared from 5 micrograms of mRNA and directionally ligated into the pBluescript II SK (+) XR vector using the pBluescript II XR cDNA Library Construction Kit according to manufacturer's instructions with modifications (Stratagene). Plasmid DNA was then transformed by electroporation into DH10B cells (Invitrogen) for propagation. Normalization was applied according to published methods (Bonaldo M.F. et al. (1996) Genome Research 6 (9):791) in order to reduce the abundance of highly expressed transcripts."

ORIGIN

Alignment Scores:	
Pred. No.:	3,448-28
Score:	290.00
Percent Similarity:	78.8%
Best Local Similarity:	65.9%
Query Match:	44.2%
DB:	7
US-10-697-787-2 (1-126)	x CV281695 (1-787)
Length:	787
Matches:	56
Conservative:	11
Mismatches:	16
Indels:	2
Gaps:	1

```

Qy      24  LeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnValAlaLysAlaValGly 43
Db      617  TTAGCCATTATGACAAAGACACTCTGACCGTTGGCATAATGTGGCTTCTGTGGTTGGC 558

Qy      44  GlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAspLeuIleAsn 63
Db      557  GGGAAATCCCGAGGAGGAGTCAAGAGGCATTATGAGATTCTGTGGAGGACTTGAATTCC 498

Qy      64  IleGluThrGlyArgValProLeuProAsnTyrLysThrPheGluSerAsnSerArgSer 83
Db      497  ATTGAAGCAGGCCAGGTGCCTTTCCCAATAATACATATCTTCGAGGCCCTAACACAGGGAG 438

Qy      84  Ile 84
Db      437  ATC 435

```

Search completed: March 10, 2006, 20:29:59
Job time : 3748 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 10, 2006, 19:14:46 ; Search time 475 Seconds
(without alignments)
1767.897 Million cell updates/sec

Title: US-10-697-787-2

Perfect score: 656

Sequence: 1 MASNGRSISPTWFSQNKMF.....HSSDFEKFQKVLVSIVSLV 126

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-Q=/abss/ABSSWEB_spool/US10697787/runat_09032006_150625_3305/app_query.fasta 1
-DB=N_Geneseq -QMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss06p
-USER=US10697787 @CGN 1 1 1096 @runat_09032006_150625_3305 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DBLEXT=7

Database :

N_Geneseq_21.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	656	100.0	381	10 ADF43565	Adf43565 Thale cre
2	473	72.1	294	14 AEA27134	Aea27134 Stress to
3	343	52.3	279	12 ADO62042	Ado62042 Transcrip
4	327.5	49.9	478	3 AAC34192	Aac34192 Arabidops

	5	290	44.2	760	13	ADR59489
	6	289	44.1	508	3	AAC35256 Arabidops
	7	287	43.8	704	13	ADX32012 Plant ful
	8	286	43.6	306	3	AAC42715 Arabidops
c	9	286	43.6	556	10	ABX57121 Arabidops
	10	286	43.6	578	10	ADD30327 Plant yie
	11	286	43.6	578	12	ADI44256 Plant tra
	12	286	43.6	578	14	AEA27034 Stress to
	13	285	43.4	482	14	ADW16587 Eucalyptu
	14	277.5	42.3	504	3	AAC49819 Arabidops
c	15	276	42.1	466	12	ADJ43786 Plant cDN
	16	275.5	42.0	294	3	AAC43000 Arabidops
	17	275.5	42.0	294	14	AEA27038 Stress to
c	18	275.5	42.0	543	13	ACN61386 Cotton gy
	19	275.5	42.0	543	13	ACN61386 Cotton gy
	20	273	41.6	303	3	AAC46353 Arabidops
	21	270.5	41.2	507	13	ADX14446 Plant ful
	22	266	40.5	636	13	ADR65153 Cotton cD
c	23	262	39.9	513	11	ACL30919 Rice abio
	24	259.5	39.6	505	13	ACN61224 Cotton gy
	25	259.5	39.6	532	13	ACN59701 Cotton gy
	26	259.5	39.6	538	13	ACN56695 Cotton gy
c	27	259.5	39.6	542	13	ACN56782 Cotton gy
	28	259.5	39.6	604	13	ADR64219 Cotton cD
	29	249	38.0	577	13	ACN59055 Cotton gy
	30	245	37.3	513	10	ADB31830 DNA encod
	31	245	37.3	513	12	ADO02248 Thalecres
	32	245	37.3	513	14	AEA27102 Stress to
	33	239	36.4	776	12	ADJ39433 Plant cDN
	34	233.5	35.6	980	13	ADX51987 Plant ful
	35	231	35.2	2192	3	AAA78533 Plant SDF
	36	228.5	34.8	584	13	ACN54517 Cotton an
	37	218	33.2	2106	14	ADW16604 Eucalyptu
	38	215	32.8	1943	14	ADW16980 Pinus rad
	39	214.5	32.7	1136	13	ADX12817 Plant ful
	40	212.5	32.4	661	12	ADO62761 Transcrip
	41	211	32.2	866	6	ABK65370 Arabidops
	42	211	32.2	866	10	ADD30950 Plant yie
	43	211	32.2	866	10	ADE37148 Plant yie
	44	211	32.2	866	12	ADI41988 Plant tra
	45	211	32.2	866	12	ADO03516 Thalecres

ALIGNMENTS

RESULT 1
ADF43565

ID ADF43565 standard; DNA; 381 BP.

XX AC ADF43565;

XX DT 12-FEB-2004 (first entry)

XX DE Thale cress peronospora parasitica resistant protein coding sequence.

XX KW transgenic plant; peronospora parasitica resistant; PPR2;
XX KW pathogen resistant phenotype; modified pathogen resistance; thale cress;
XX KW gene; ds.

XX OS Arabidopsis thaliana.

XX FH Key Location/Qualifiers

XX FT CDS 1..381

XX FT /*tag= a

XX FT /product= "Thale cress peronospora parasitica resistant protein"

XX PN WO2003091412-A2.

XX PD 06-NOV-2003.

XX XX 24-APR-2003; 2003WO-US012981.

```
PR 24-APR-2002; 2002US-0375333P.
XX
XX (AGRI-) AGRINOMICS LLC.
PA
XX Federapfel N, Lammers A, Liu XL, Bates SR, Westerlund C;
PI Fitch JR;
PI
XX
XX WPI; 2003-865582/80.
DR P-PSDB; ADF43566.
XX
XX New transgenic plants with increased resistance to pathogens due to
PT altered expression of Peronospora Parasitica Resistant gene (PPR2),
PT useful for generating plants with a pathogen resistance phenotype.
XX
XX Example 4; SEQ ID NO 1; 36pp; English.
XX
XX The invention relates to a transgenic plant that possesses a plant
CC transformation vector comprising a nucleotide sequence that encodes a
CC peronospora parasitica resistant (PPR2) protein, or a PPR2 orthologue.
CC The transgenic plant is useful in generating plants with a pathogen
CC resistance phenotype. The PPR2 nucleic acids and proteins of the
CC invention are useful in the generation of genetically modified plants
CC having a modified pathogen resistance phenotype. The present DNA sequence
CC encodes an Arabidopsis thaliana PPR2 protein of the invention.
XX
XX Sequence 381 BP; 122 A; 76 C; 76 G; 107 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 4.3e-76 Length: 381
Score: 656.00 Matches: 126
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0
US-10-697-787-2 (1-126) x ADF43565 (1-381)
Qy 1 MetAlaSerAsnSerArgSerIleSerProTyrThrPheSerGlnAsnLysMetPhe 20
Db 1 ATGGCGTCAAACTCAAGAGTTCAATCTCACCATTGGAGCTTTAGTCAAAACAAGATGTTTC 60
Qy 21 GluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnValAlaLys 40
Db 61 GAGAGGGCCCTTGGCAGTTTACGACAGGACACACCCGCCGATGGCAATGTGGCAAAA 120
Qy 41 AlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAsp 60
Db 121 GCTGTGCGAGGAGAAACTGTAGAGAGTGAAGGCCCACTATGCATCTTCGTCGAGGAT 180
Qy 61 LeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrIleThrPheGluSerAsn 80
Db 181 CTCATCAACATCGAGACTGGTGTGTCCTTTGCCCAATTACAGAGCTTTCGAATCTTAAC 240
Qy 81 SerArgSerIleAsnAspPheAspThrArgTyrIleThrLysTyrLeuTyrMetLeu 100
Db 241 TCAAGAAGCATCAATGACTTTGACACAGGTATATACTAAATATCTATATATGATGCTC 300
Qy 101 SerIleTyrPheAspAsnHisSerSerAspPheGluLysPheSerGlnLysValLeuVal 120
Db 301 TCGATATATTTTGATATATCACTTAGTGATTTTGAGAAATCTCTCAAAAGTTCTTGTA 360
Qy 121 SerTyrIleSerLeuVal 126
Db 361 AGTTATATTTCTTTGGTT 378
RESULT 2
AEA271134
ID AEA271134 standard; DNA; 294 BP.
XX
XX AEA271134;
XX
XX 28-JUL-2005 (first entry)
XX
```

```
DE Stress tolerant plant-related transcription factor gene SeqID975.
XX
XX transcription factor; transgenic plant; agriculture; drought resistance;
KW stress tolerance; gene; ds.
XX
XX Arabidopsis thaliana.
OS
XX
XX WO2005047516-A2.
PN
XX
XX 26-MAY-2005.
PD
XX
XX 12-NOV-2004; 2004WO-US037584.
PF
XX
XX 13-NOV-2003; 2003US-00714887.
PR
XX 05-DEC-2003; 2003US-0527658P.
PR
XX 05-FEB-2004; 2004US-0542928P.
PR
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
PA
XX Heard JE, Riechmann JL, Creelman RA, Ratcliffe OJ, Canales RD;
PI Repetti P, Kumimoto RW, Gutterman NI, Reuber TL, Pineda O;
PI Sherman BK, Morrison TA, Keddie JS, Jiang C, Century KS, Adam L;
PI Zhang JZ, Hempel FD, Libby JM;
PI
XX WPI; 2005-372386/38.
DR P-PSDB; AEA271135.
DR
XX
XX New transgenic plants for producing commercially or agriculturally useful
PT plants having improved tolerance to drought, shade and low nitrogen
PT conditions.
PT
XX Disclosure; SEQ ID NO 975; 407pp; English.
PS
XX This invention relates to a novel plant transcription factor
XX polypeptides, the DNA sequences which encode them and their use in
XX creating transgenic plants. The transgenic plant and methods are useful
XX for producing commercially or agriculturally useful plants having
XX improved tolerance to drought, shade and low nitrogen conditions when
XX compared to wild-type reference plants. The present sequence is that of
XX a plant transcription factor gene which was used during the development of
XX the transgenic plants of the invention.
XX
XX Sequence 294 BP; 99 A; 65 C; 67 G; 63 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 2.39e-52 Length: 294
Score: 473.00 Matches: 90
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 72.1% Indels: 0
DB: 14 Gaps: 0
US-10-697-787-2 (1-126) x AEA271134 (1-294)
Qy 1 MetAlaSerAsnSerArgSerIleSerProTyrThrPheSerGlnAsnLysMetPhe 20
Db 1 ATGGCGTCAAACTCAAGAGTTCAATCTCACCATTGGAGCTTTAGTCAAAACAAGATGTTTC 60
Qy 21 GluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnValAlaLys 40
Db 61 GAGAGGGCCCTTGGCAGTTTACGACAGGACACACCCGCCGATGGCAATGTGGCAAAA 120
Qy 41 AlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAsp 60
Db 121 GCTGTGCGAGGAGAAACTGTAGAGAGTGAAGGCCCACTATGCATCTTCGTCGAGGAT 180
Qy 61 LeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrIleThrPheGluSerAsn 80
Db 181 CTCATCAACATCGAGACTGGTGTGTCCTTTGCCCAATTACAGAGCTTTCGAATCTTAAC 240
Qy 81 SerArgSerIleAsnAspPheAspThrArg 90
Db 241 TCAAGAAGCATCAATGACTTTTGACACAAGG 270
```

CC content, altered seed protein content, altered seedprenyl content,
CC altered leaf prenyl lipid content, increased anthocyanin levels, and
CC decreased anthocyanin levels. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 279 BP; 86 A; 56 C; 65 G; 72 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,11e-35 Length: 279
Score: 343.00 Matches: 69
Percent Similarity: 80.4% Conservative: 5
Best Local Similarity: 75.0% Mismatches: 16
Query Match: 52.3% Indels: 2
DB: 12 Gaps: 1

US-10-697-787-2 (1-126) x ADO62042 (1-279)

QY 1 MetAlaSerAsnSerArgSerSerIleSerProTTPThrPheSerGlnAsnLysMetPhe 20
DB 1 ATGGCTCTAGTCTCTATGAGCTCGAGCTCTTCTTGACGCTCTAAGCAAAACGATGTTTC 60
QY 21 GluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTTPHisAsnValAlaLys 40
DB 61 GAGAGGGCTTTAGCCGTTTACGATTAAGACACTCCGACCGTTGGCAAAACGTTGCTAA 120
QY 41 AlaValGlyGlyThrValGluGluValLysArgHisTyrAspIleLeuValGluAsp 60
DB 121 GCAGTTGGAAAGTAAATCTGCAGAGGAAGTTAAACGCTCACTAGCAGATCTCGTTGAAGAT 180
QY 61 LeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGlu----- 78
DB 181 CTCATGAACATCGAACAAAGACTTAGTACTCTTGGCTTAATAACAAACCGTCGATGTGGA 240
QY 79 SerAsnSerArgSerIleAsnAspPheAspThrArg 90
DB 241 AGTAAATCTAGAGGCGATCGATGATTTCGATTGAGG 276

RESULT 4
AAC34192
ID AAC34192 standard; DNA; 478 BP.
XX
AC AAC34192;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 5783.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.

RESULT 3
ADO62042
ID ADO62042 standard; DNA; 279 BP.
XX
AC ADO62042;
XX
DT 15-JUL-2004 (first entry)
XX
DE Transcription factor G2723 coding sequence, SEQ ID 509.
XX
KW Plant; transcription factor; transgenic plant; abiotic stress tolerance;
KW osmotic stress tolerance; cold tolerance; heat tolerance;
KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
KW glyphosate resistance; flowering; fertility; seed development; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO2004031349-A2.
XX
XX 15-APR-2004.
XX
PF 18-SEP-2003; 2003WO-US030292.
XX
PR 18-SEP-2002; 2002US-0411837P.
PR 17-DEC-2002; 2002US-0434166P.
PR 24-APR-2003; 2003US-0465809P.
XX
PA (MEND-) MENDEL BIOTECHNOLOGY INC.
XX
XX Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;
PI Riechmann JL, Haake V, Dubell AN, Keddle JS, Sherman BK;
XX
DR WPI: 2004-330163/30.
DR P-PSDB; ADO62043.
XX
XX New recombinant polynucleotide encoding transcription factor
PT polypeptides, useful for producing transgenic plants with advantageous
PT properties compared to a reference plant.
XX
PS Claim 1; SEQ ID NO 509; 510pp; English.
XX
XX The present invention relates to novel plant transcription factor
CC proteins (I) and nucleotide sequences (II) (ADO61534-ADO63778). The
CC sequences can be used to produce transgenic plants, which overexpress
CC (II), where the transgenic plant has an altered trait as compared to a
CC non-transgenic plant or wild-type plant. The transgenic plant comprises
CC an altered trait selected from increased tolerance to abiotic stress,
CC increased tolerance to osmotic stress, increased tolerance to cold,
CC increased germination in cold, increased tolerance to heat, increased
CC germination in heat, increased tolerance to freezing conditions,
CC increased tolerance to low nitrogen conditions, increased tolerance to
CC low phosphate conditions, increased tolerance to disease, including
CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,
CC increased tolerance to multiple fungal pathogens, increased resistance to
CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,
CC increased sensitivity to ACC, altered sugar sensing, increased tolerance
CC to sugars, altered carbon/nitrogen sensing, early flowering, late
CC flowering, altered flower structure, loss of flower determinacy, reduced
CC fertility, altered shoot meristem development, altered branching pattern,
CC altered stem morphology, altered vascular tissue structure, reduced
CC apical dominance, altered trichome density, altered trichome development,
CC altered trichome structure, altered root development, altered shade
CC avoidance, altered seed development, altered seed ripening, altered seed
CC germination, slow growth, fast growth, altered cell differentiation,
CC altered cell proliferation, altered cell expansion, altered phase change,
CC altered senescence, abnormal embryo development, altered programmed cell
CC death, lethality when overexpressed, altered necrosis patterns, increased
CC plant size, increased biomass, large seedlings, dwarfed plants, dark
CC green leaves, change in leaf shape, increased leaf size and mass, light
CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
CC altered seed coloration, altered seed size, altered seed shape, large
CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil

PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.

PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.

PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161320P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.: 4,64e-33 Length: 478
Score: 327.50 Matches: 67
Percent Similarity: 73.2% Conservative: 4
Best Local Similarity: 69.1% Mismatches: 19
Query Match: 49.9% Indels: 7
DB: 3 Gaps: 2

US-10-697-787-2 (1-126) x AAC34192 (1-478)

QY 10 SerProTrrPheSerGlnAsnLysMetPheGluAtcAlaLeuAlaValTrrPheLys 29
DB 5 TCTTCTTGACGCTTAAGCAAAACAGATGTTGAGAGGCGTTTACGCTTTACGATAAA 64
QY 30 AspThrProAspArgTrrPheAsnValAlaLysAlaValGlyGlyLysThrValGluGlu 49
DB 65 GACATCCCGACCGTTGGCAAAACGTCGCTAAGCAGTTGGAAGTAATCTCCAGAGNA 124
QY 50 ValLysArgHisTrrPheAspLeuValGluAspLeuLeuAsnLeuThrGlyArgVal 69
DB 125 GTTAAACGTCCTACGACATCTCGTTGAAGATCTCATGAACATCGAACAGACTTAGTA 184
QY 70 ProLeuProAsnTrrLysThrPheGlu-----SerAsnSerArgSerIleAsnAspPhe 87
DB 185 CTTTTCCTTAAATACAAACCGTCGATGTTGGAAGTAATCTAGAGGAATCAATGATTTT 244
QY 88 AspThrArgTrrIleThrLysTrr-----LeuTrrMetMet 99
DB 245 GATTTGAGGTTATTGAGGAATATGAGATCCAGTGAAGCTGTACATGATG 295

RESULT 5

ADR59489
ID ADR59489 standard; cDNA; 760 BP.

AC ADR59489;

DT 02-DEC-2004 (first entry)

DE Cotton cDNA sequence, SEQ ID 270.

KW Cotton; ss; plant; cold tolerance; growth rate; cell cycle pathway;
KW drought tolerance; plant disease resistance; galactomannan; lignin;
KW plant growth regulator; heat tolerance; herbicide tolerance;
KW homologous recombination; extreme osmotic condition tolerance;
KW pathogen resistance; pest resistance; yield; photosynthesis; seed oil;
KW stress resistance.

OS Gossypium hirsutum.

XX US2004181830-A1.

PN

XX 16-SBP-2004.
XX 29-JAN-2004; 2004US-00767795.
XX 07-MAY-2001; 2001US-00849529.
PR 12-DEC-2001; 2001US-00021323.
XX (KOVA/) KOVALIC D K.
PA (ZHOU/) ZHOU Y.
PA (CAOY/) CAO Y.
XX Kovalic DK, Zhou Y, Cao Y;
DR WPI; 2004-667718/65.
XX New recombinant nucleic acid molecules and polypeptides from Gossypium
PT hirsutum, useful for producing plants with improved biological
PT characteristics (e.g. improved plant cold or drought tolerance).
XX Claim 1; SEQ ID NO 270; 14pp; English.

CC The invention relates to a recombinant polynucleotide comprising any of
CC the 58798 Cotton plant cDNA sequences mentioned in the specification.
CC Also a recombinant polypeptide comprising any of the 58798 amino acid
CC sequences mentioned in the specification and producing a plant having an
CC improved property. Producing a plant having an improved property
CC comprises transforming a plant with a recombinant construct comprising a
CC promoter region functional in a plant cell operably joined to a
CC polynucleotide comprising a coding sequence for a polypeptide associated
CC with the property, and growing the transformed plant. The polypeptide is
CC useful for improving plant cold tolerance, manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, improving plant
CC drought tolerance, providing increased resistance to plant disease,
CC producing galactomannan (or lignin or plant growth regulators), improving
CC plant heat tolerance, improving plant tolerance to herbicides, increasing
CC the rate of homologous recombination in plants, improving plant tolerance
CC to extreme osmotic conditions or to pathogens or pests, improving yield
CC and/or content, improving yield by modification of carbohydrate, nitrogen
CC or phosphorus use and/or uptake, or improving yield by providing improved
CC plant growth and development under at least one stress condition. The
CC polynucleotide and polypeptide may also be used in recombinant DNA
CC constructs, in physical arrays of molecules, as plant breeding markers,
CC or in computer-based storage and analysis systems. The present sequence
CC is a Cotton plant cDNA of the invention. NOTE: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20040181830. However only 6585
CC polynucleotide sequences were available, the remaining 52213
CC polynucleotides and all 58798 protein sequences were not present.

SQ Sequence 760 BP; 308 A; 90 C; 144 G; 218 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6.9e-28 Length: 760
Score: 290.00 Matches: 56
Percent Similarity: 83.8% Conservative: 11
Best Local Similarity: 70.0% Mismatches: 11
Query Match: 44.2% Indels: 2
DB: 13 Gaps: 1

US-10-697-787-2 (1-126) x ADR59489 (1-760)

QY 1 MetAlaSerAsnSerArgSerSerIleSerProTrrPheSerGlnAsnLysMetPhe 20
DB 113 ATGTCATCGAATTCATGTCT-----GCTTCATGGACAGCAACAAACAGATTTC 166
QY 21 GluArgAlaLeuAlaValTrrAspLysAspThrProAspArgTrrPheAsnValLys 40
DB 167 GAAAGGGCTTTAGCTGTTTACGCAAGGACACACAGATCGTTGGTACAAATGTTCTAAA 226
QY 41 AlaValGlyGlyLysThrValGluGluValLysArgHisTrrAspIleLeuValGluAsp 60

[illegible]

CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.

XX
SQ Sequence 704 BP; 202 A; 159 C; 121 G; 222 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.53e-27 Length: 704
Score: 287.00 Matches: 54
Percent Similarity: 80.2% Conservative: 11
Best Local Similarity: 66.7% Mismatches: 16
Query Match: 43.8% Indels: 0
DB: 13 Gaps: 0

US-10-697-787-2 (1-126) x ADX32012 (1-704)

Qy 1 MetAlaSerAsnSerArgSerIleSerProTrpThrPheSerGlnAsnLysMetPhe 20
Db 150 ATGGCATCCAGTTCAATCTCAGCCTCTGGCTCATGGAGTAAAGGACAAAGGCGCTTT 209
Qy 21 GluArgAlaLeuAlaValTyAspLysAspThrProAspArgTrpHisAsnValAlaLys 40
Db 210 GAAAGGCTTAGCTGTTATGACNAGGACACTCTGACCGTTGGTACATGTGTCAT 269
Qy 41 AlaValGlyGlyThrValGluGluValLysArgHisTyAspIleLeuValGluAsp 60
Db 270 GCTGTTGGTGCNAAACTCCAGAGGAAGTGAAGAGGCACTACGAACTCTTGTTCAGCAT 329
Qy 61 LeuIleAsnIleGluThrGlyArgValProLeuProAsnTyIysThrPheGluSerAsn 80
Db 330 GTTAAGCATATTGAGTCTGGACGTGTGCCATTCCNAAATTACAGAAAACTACTTTCAGGG 389
Qy 81 Ser 81
Db 390 TCA 392

RESULT 8

AAC42715

ID AAC42715 standard; DNA; 306 BP.

XX AAC42715;

AC AAC42715;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 36586.

XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

XX 23-MAR-1999; 99US-0123548P.

XX 29-MAR-1999; 99US-0126264P.

XX 01-APR-1999; 99US-0127462P.

XX 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0128845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.

PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-01443352P.
PR 20-JUL-1999; 99US-0144332P.
PR 20-JUL-1999; 99US-0144384P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145313P.
PR 27-JUL-1999; 99US-0145318P.
PR 27-JUL-1999; 99US-0145319P.
PR 28-AUG-1999; 99US-0145351P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0148684P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156598P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.

PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 21-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160988P.
PR 25-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.: 6.63e-28 Length: 306
Score: 286.00 Matches: 56
Percent Similarity: 74.7% Conservative: 6
Best Local Similarity: 67.5% Mismatches: 19
Query Match: 43.6% Indels: 2
DB: 3 Gaps: 1

US-10-697-787-2 (1-126) x AAC42715 (1-306)

QY 1 MetAlaSerAsnSerArgSerSerile-----SerProTrpThrPheSerGlnAsnLys 18
DB 1 ATGGCATCAGGCTCAATGCTTCATTATGGCTCTGGCTCATGGAGCTGTTAAGCAGACAAA 60
QY 19 MetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnVal 38
DB 61 GCCTTTGAGCGTCTAGCAGTCTATGCCAAGACACTCCGACCGTGGCACAATGTT 120
QY 39 AlaLysAlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleuVal 58
DB 121 GCTAGAGCTGTTGGTGTAAACACACAGAAAGCTAAGACAGACAGTATGACCTTCTAGTT 180
QY 59 GluAspLeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrIleThrPheGlu 78
DB 181 CGTGACATCGAAGCATCGAAGTGGTCAGTGGCCATTCCTGTGACTACAAGACTACTACA 240
QY 79 SerAsnSer 81
DB 241 GGAAACAGC 249

RESULT 9

ABX57121/c
ID ABX57121 standard; DNA; 556 BP.
XX
AC ABX57121;
XX
DT 20-FEB-2003 (first entry)
XX
DE Arabidopsis thaliana polynucleotide #473.
XX
KW Thale cress; gene; ds; genetic manipulation; plant; biosynthesis;
KW genetic modification; environmental stress; disease resistance;
KW fungicide; insecticide; stress tolerance.
XX
OS Arabidopsis thaliana.
XX

PN US2002040489-A1.
XX 04-APR-2002.
PD 26-JAN-2001; 2001US-00770152.
XX 27-JAN-2000; 2000US-0178503P.
PR
XX (GORL/) GORLACH J.
PA (ANY/) AN Y.
PA (HAMI/) HAMILTON C M.
PA (PRIC/) PRICE J L.
PA (RAIN/) RAINES T M.
PA (YUY/) YU Y.
PA (RAME/) RAMEAKA J G.
PA (PAGE/) PAGE A.
PA (MATH/) MATHW A V.
PA (LEDF/) LEDFORD B L.
PA (WOES/) WOESSNER J P.
PA (HAAS/) HAAS W D.
PA (GARC/) GARCIA C A.
PA (KRIC/) KRICKER M.
PA (SLAT/) SLATER T.
PA (DAVI/) DAVIS K R.
PA (ALLE/) ALLEN K.
PA (HOFF/) HOFFMAN N.
PA (HURB/) HURBAN P.
XX
PI Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
PI Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;
PI Hurban P;
XX
DR WPI; 2003-110410/10.
XX
PT Novel Arabidopsis thaliana nucleic acid useful for identifying homologous
PT or related genes, and to create genetically modified and transgenic
PT organisms, such as plant cells and plants.
XX
PS Claim 1; SEQ ID NO 473; 45pp; English.
XX
CC The invention relates to Arabidopsis thaliana nucleic acid sequences. The
CC DNA sequences and the polypeptides they encode are useful for identifying
CC homologous or related genes, for producing compositions that modulate the
CC expression or function of the polypeptides, for mapping functional
CC regions of the protein, in diagnosis, for studying associated
CC physiological pathways, for genetic manipulation of cells, preferably
CC plant cells, in screening assays of various plant strains to determine
CC the strains that are capable of withstanding a particular disease or
CC environmental stress, for enhancing or inhibiting production of
CC biosynthetic products in plants and to create genetically modified and
CC transgenic organisms, such as plant cells and plants. Transgenic plants
CC are useful for introducing or improving disease resistance and stress
CC tolerance in plants, screening biologically active agents, such as
CC fungicides and insecticides, and for elucidating biochemical pathways.
CC Sequences ABX5649-ABX57647 represent Arabidopsis thaliana
CC polynucleotides of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification but was obtained in
CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html
XX
SQ Sequence 556 BP; 146 A; 114 C; 115 G; 181 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.5e-27 Length: 556
Score: 286.00 Matches: 56
Percent Similarity: 74.7% Conservative: 6
Best Local Similarity: 67.5% Mismatches: 19
Query Match: 43.6% Indels: 2
DB: 10 Gaps: 1

US-10-697-787-2 (1-126) x ABX57121 (1-556)

Qy 1 MetAlaSerAsnSerArgSerSerIle-----SerProTrpThrPheSerGlnAsnLys 18

Db 474 ATGGCATCAGGCTCAATGCTCTTATGGCTCTGGCTCATGGACTGTTAAGCAGACAAA 415
Qy 19 MetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnVal 38
Db 414 GCCTTTGAGGGTCTCTAGCAGTCTATGACCAAGACACTCGGACCGTTGGCACAATGTT 355
Qy 39 AlaLysAlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuVal 58
Db 354 GCTAGAGCTGTTGGTGGTAAACACCCAGAAGAAGCTTAAGAGACAGTATGACCTTCTAGTT 295
Qy 59 GluAspLeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGlu 78
Db 294 CGTGACATCAAGACATCGAAGATGTCAGTGCCATTCCTGACTACAAGACTACTACA 235
Qy 79 SerAsnSer 81
Db 234 GGAACACAGC 226

RESULT 10
ADD30327
ID ADD30327 standard; cDNA; 578 BP.
XX
AC ADD30327;
XX
DT 15-JAN-2004 (first entry)
XX
DE Plant yield-related polynucleotide clone G1789.
XX
KW ds; transcription factor; transgenic plant; growth rate; senescence;
KW seed germination rate; plant vigor; seedling vigor.
XX
OS Arabidopsis thaliana.
XX
PN WO2003013227-A2.
XX
PD 20-FEB-2003.
XX
PF 09-AUG-2002; 2002WO-US025805.
XX
PR 09-AUG-2001; 2001US-0310847P.
PR 19-NOV-2001; 2001US-0336049P.
PR 11-DEC-2001; 2001US-0338692P.
PR 14-JUN-2002; 2002US-00171468.
XX
PA (MEND-) MENDEL BIOTECHNOLOGY INC.
XX
PI Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE;
PI Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;
PI Broun PE;
XX
DR WPI; 2003-248221/24.
DR P-PSDB; ADD30328.
XX
PT New plant transcription factor polynucleotides and polypeptides, useful
PT in producing transgenic plants with commercially valuable properties
PT such as an alteration in a plant growth characteristic, e.g. growth rate
PT or apomixis.
XX
PS Disclosure; SEQ ID NO 356; 454pp; English.
XX
CC The invention relates to a number of isolated Arabidopsis thaliana cDNA
CC sequences and their encoded proteins which are especially transcrip
CC tion factor related cDNA's and proteins. The isolated or recombinant plant
CC transcription factor polynucleotides and polypeptides are useful in
CC producing transgenic plants with commercially valuable properties, i.e.
CC modified or altered desirable traits as compared to a reference plant,
CC such as an alteration in a plant growth characteristic, e.g. growth rate,
CC germination rate of seeds, vigor of plants and seedlings, or leaf and
CC flower senescence. Sequence information related to the polynucleotides
CC and polypeptides can also be used in bioinformatic search methods. The
CC transgenic plant is useful for growing a progeny plant from a parent
CC plant. This sequence represents one of the cDNAs of the invention.

```
XX SQ Sequence 578 BP; 188 A; 119 C; 117 G; 154 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.58e-27 Length: 578
Score: 286.00 Matches: 56
Percent Similarity: 74.7% Conservative: 6
Best Local Similarity: 67.5% Mismatches: 19
Query Match: 43.6% Indels: 2
DB: 10 Gaps: 1

US-10-697-787-2 (1-126) x ADD30327 (1-578)
QY 1 MetAlaSerAsnSerArgSerIle-----SerProTrrPrrPheSerGlnAsnLys 18
DB 108 ATGGCATCAGGCTCAATGCTTCTTATGGCTCTGGCTCATGGACTGTTAAGCAGACAAA 167
QY 19 MetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnVal 38
DB 168 GCCTTTGAGCGTGTCTAGCAGTCTATGACCAAGACACTCCGACCGTTGGCACAATGTT 227
QY 39 AlaLysAlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuVal 58
DB 228 GCTAGAGCTGTTGGTGTAAACACACAGAAAGCTAAGAGACAGTATGACCTTCTAGTT 287
QY 59 GluAspLeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGlu 78
DB 288 CGTGACATCGAAAGCATCGAAGTGTACGTGCGCATTCCTGACTACAGACTACTACA 347
QY 79 SerAsnSer 81
DB 348 GGAACACAGC 356

RESULT 11
ADI44256
ID ADI44256 standard; DNA; 578 BP.
XX AC ADI44256;
XX DT 22-APR-2004 (first entry)
XX DE Plant transcription factor related polynucleotide #1717.
XX KW transgenic; plant; enhanced tolerance to abiotic stress;
KW glyophosphate tolerance; hormone sensitivity; disease resistance;
KW sugar sensing; flowering; flower structure; stem bifurcation;
KW branching pattern; apical dominance; trichome; stem morphology;
KW root growth; root hair; seed development; cell proliferation;
KW cell differentiation; premature senescence; necrosis; plant size;
KW leaf morphology; seed morphology; seed biochemistry; root anthocyanin;
KW plant anthocyanin; light response; shade avoidance; bioinformatic;
KW transcription factor; ds.
XX OS Unidentified.
XX PN US2004019927-A1.
XX PD 29-JAN-2004.
XX PF 25-FEB-2003; 2003US-00374780.
XX PR 18-APR-2001; 2001US-00837944.
XX PA (SHER/) SHERMAN B K.
XX PA (RIE/) RIECHMANN J L.
XX PA (JIAN/) JIANG C.
XX PA (HEAR/) HEARD J E.
XX PA (HAKE/) HAAKE V.
XX PA (CREE/) CREELMAN R A.
XX PA (RATC/) RATCLIFFE O.
XX PA (ADAM/) ADAM L J.
XX PA (REUB/) REUBER T L.
XX PA (KEDD/) KEDDIE J.
```

```
PA (BROU/) BROUN P E.
PA (PILG/) PILGRIM M L.
PA (DUBE/) DUBELL A N.
PA (PINE/) PINEDA O.
PA (YUGG/) YU G.
XX Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V,
PI Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddie J, Broun PB;
PI Pilgrim ML, Dubell AN, Pineda O, Yu G;
XX WPI; 2004-132245/13.
DR P-PSDB; ADI44257.
XX New transgenic plant comprising a recombinant polynucleotide of any one
PT of more than 500 nucleotide sequences, useful in bioinformatic search
PT methods.
XX Disclosure; SEQ ID NO 2719; 435pp; English.
CC The invention describes a transgenic plant comprising a recombinant
CC polynucleotide of any one of more than 500 nucleotide sequences fully
CC defined in the specification or its complement. The method of the
CC invention can be used to produce a plant having altered traits such as:
CC enhanced tolerance to abiotic stress; glyophosphate tolerance; hormone
CC sensitivity; disease resistance; sugar sensing; early or late flowering;
CC altered flower structure, change in stem bifurcations, altered branching
CC pattern, reduced apical dominance, reduced trichome density; lack of
CC trichomes; reduced ectopic trichome development; altered trichome
CC development; increase in trichome number; altered stem morphology;
CC increased root growth; increased root hairs; altered seed development;
CC altered cell proliferation or cell differentiation; rapid development;
CC premature senescence; increased necrosis; increase in seedling or plant
CC size; decreased plant size; leaf morphology; seed morphology; seed
CC biochemistry; increase in root anthocyanins; increase in plant
CC anthocyanins, or alteration in light response or shade avoidance. The
CC transgenic plant, polynucleotides and polypeptides are useful in
CC bioinformatic search methods. This sequence represents a plant
CC transcription factor related polynucleotide.
XX SQ Sequence 578 BP; 188 A; 119 C; 117 G; 154 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 1.58e-27 Length: 578
Score: 286.00 Matches: 56
Percent Similarity: 74.7% Conservative: 6
Best Local Similarity: 67.5% Mismatches: 19
Query Match: 43.6% Indels: 2
DB: 10 Gaps: 1

US-10-697-787-2 (1-126) x ADI44256 (1-578)
QY 1 MetAlaSerAsnSerArgSerIle-----SerProTrrPrrPheSerGlnAsnLys 18
DB 108 ATGGCATCAGGCTCAATGCTTCTTATGGCTCTGGCTCATGGACTGTTAAGCAGACAAA 167
QY 19 MetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnVal 38
DB 168 GCCTTTGAGCGTGTCTAGCAGTCTATGACCAAGACACTCCGACCGTTGGCACAATGTT 227
QY 39 AlaLysAlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuVal 58
DB 228 GCTAGAGCTGTTGGTGTAAACACACAGAAAGCTAAGAGACAGTATGACCTTCTAGTT 287
QY 59 GluAspLeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGlu 78
DB 288 CGTGACATCGAAAGCATCGAAGTGTACGTGCGCATTCCTGACTACAGACTACTACA 347
QY 79 SerAsnSer 81
DB 348 GGAACACAGC 356

RESULT 12
AEA27034
```

ID AEA27034 standard; DNA; 578 BP.
XX AC AEA27034;
XX DT 28-JUL-2005 (first entry)
XX DE Stress tolerant plant-related transcription factor gene SeqID875.
XX KW transcription factor; transgenic plant; agriculture; drought resistance;
XX KW stress tolerance; gene; da.
XX KW Arabidopsis thaliana.
XX OS
XX PN WO2005047516-A2.
XX PD 26-MAY-2005.
XX PF 12-NOV-2004; 2004WO-US037584.
XX PR 13-NOV-2003; 2003US-00714887.
XX PR 05-DEC-2003; 2003US-0527658P.
XX PR 05-FEB-2004; 2004US-0542928P.
XX PA (WEND-) MENDEL BIOTECHNOLOGY INC.
XX PI Heard JE, Riechmann JL, Creelman RA, Ratcliffe OJ, Canales RD;
PI Repetti P, Kumimoto RW, Guttererson NI, Reuber TL, Pineda O;
PI Sherman BK, Morrison TA, Keddle JS, Jiang C, Century KS, Adam L;
PI Zhang JZ, Hempel PD, Libby JM;
XX PS WPI; 2005-372386/38.
XX DR P-PSDB; AEA27035.
XX PT New transgenic plants for producing commercially or agriculturally useful
XX PT plants having improved tolerance to drought, shade and low nitrogen
XX PT conditions.
XX PS Disclosure; SEQ ID NO 875; 407pp; English.
XX CC This invention relates to a novel plant transcription factor
XX CC polypeptides, the DNA sequences which encode them and their use in
XX CC creating transgenic plants. The transgenic plant and methods are useful
XX CC for producing commercially or agriculturally useful plants having
XX CC improved tolerance to drought, shade and low nitrogen conditions when
XX CC compared to wild-type reference plants. The present sequence is that of a
XX CC plant transcription factor gene which was used during the development of
XX CC the transgenic plants of the invention.
XX SQ Sequence 578 BP; 188 A; 119 C; 117 G; 154 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.59e-27 Length: 578
Score: 286.00 Matches: 56
Percent Similarity: 74.7% Conservative: 6
Best Local Similarity: 67.5% Mismatches: 19
Query Match: 43.6% Indels: 2
DB: 14 Gaps: 1

US-10-697-787-2 (1-126) x AEA27034 (1-578)

Qy 1 MetAlaSerAenSerArgSerSerIle-----SerProTtrPthPheSerGlnAenLys 18
Dy 108 ATGCATCAGGCTCAATGTCTTCTTATGGCTCTGGCTCATGGACTGTTAAGCAGAACAAA 167
Qy 19 MetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTTPHisAenVal 38
Dy 168 GCCTTTGAGGTGCTCTAGCAGCTATGACACACACTCCGGACCTTGGCACAATGTT 227
Qy 39 AlalysAlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuVal 58
Dy 228 GCTAGACCTGTGTGGTGTAAACACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 287
Qy 59 GluAspLeuIleAenIleGluThrGlyArgValProLeuProAenTyrLysThrPheGlu 78

Db 288 CGTGACATCGAAAGCATCGAGAAATGTCACGTGCCATTCCTGACTACAGACTACTACA 347
Qy 79 SerAenSer 81
Dy 348 GGAAACAGC 356
RESULT 13
ADW16587
ID ADW16587 standard; cDNA; 482 BP.
XX AC ADW16587;
XX DT 24-MAR-2005 (first entry)
XX DE Eucalyptus grandis transcription factor cDNA MYB family Seq 306.
XX KW gene; ss; plant; transcription; gene regulation; gene expression;
XX KW transgenic plant; drought resistance; disease resistance; salt tolerance;
XX KW cold tolerance; freezing tolerance; flowering; flavor enhancer;
XX KW flower color.
XX OS Eucalyptus grandis.
XX PN WO2005001050-A2.
XX PD 06-JAN-2005.
XX PF 07-JUN-2004; 2004WO-US017965.
XX PR 06-JUN-2003; 2003US-0476189P.
XX PA (ARBO-) ARBORGEN LLC.
XX PI Blokeberg LM, Bryant C, Connett MB, Emerson SJ, Frost MJ;
PI Forster RLS, Grigor M, Higgins C, Lasham A, Lund ST, Magusin A;
PI Phillips J, Puthigae S, Veerakone S, Westwood C, Gause K, Wood M;
XX PS WPI; 2005-075542/08.
XX DR P-PSDB; ADW17373.
XX CC New polynucleotides isolated from plants encoding transcription factors,
XX CC and polypeptides encoded by such polynucleotides, useful for regulating
XX CC gene transcription and gene expression.
XX PS Claim 3; SEQ ID NO 306; 1265pp; English.
XX CC This invention relates to novel isolated plant nucleic acid molecules, or
XX CC variants thereof, that encode transcription factors. Specifically, it
XX CC refers to transcription factor proteins that are capable of binding to
XX CC DNA in order to regulate gene transcription and gene expression in a
XX CC plants, in particular Eucalyptus grandis and Pinus radiata. The present
XX CC invention describes DNA constructs containing DNA encoding a
XX CC transcription factor that regulates the promoter, which is operably
XX CC linked to the desired nucleic acid to be expressed. It further provides
XX CC transgenic plants expressing a transcription factor that confers a trait
XX CC to the plant such as increased drought, salt or disease tolerance, height
XX CC change, enhanced cold/ frost tolerance, enhanced color, health and
XX CC nutritional characteristics, as well as improved taste, starch
XX CC composition, flower longevity and germination, amongst others.
XX CC Accordingly, such plants that are successfully transfected with a DNA
XX CC construct can be characterized by a difference in flower color, petal or
XX CC leaf shape and size, aroma or plant height. This polynucleotide is a
XX CC plant transcription factor cDNA sequence of the invention.
XX SQ Sequence 482 BP; 137 A; 89 C; 121 G; 135 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.67e-27 Length: 482
Score: 285.00 Matches: 55
Percent Similarity: 84.0% Conservative: 13
Best Local Similarity: 67.9% Mismatches: 11
Query Match: 43.4% Indels: 2

DB:	14	Gaps:	2
US-10-697-787-2 (1-126) x ADM16587 (1-482)			
Qy	1	MetAlaSerAenSerArGSerSerIleSerProTrrPThrPheSerGlnAenIlyeMetPhe	20
Db	40	ATGCGATCGAATTTCTTGACTTCT---TCATCTCTGGACGCCGGAAGCAGACAAGATGTTCC	96
Qy	21	GluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrrPHisAenValAlaLys	40
Db	97	GAGAAAGGCAATGGCTCAATATGACAAGCACATCCCGACAGGTGGCAGAAGATTCGCCAAG	156
Qy	41	AlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAsp	60
Db	157	GCCCTGGGTGGGAAATCTGCAGATGAAGTGAAGACACTATGAAATTTTAAATCGAGGAC	216
Qy	61	LeuIleAenIleGluThrGlyArgValProLeuProAenTyrLysThrPheGluSerAen	80
Db	217	GTCAAGCACATCGAGTCTCGCAGAGTTCCTTTTCTTAATTACAGGTGCG---AGCAACAAT	273
Qy	81	Ser	81
Db	274	AGC	276
RESULT 14			
ID	AAC49819		
AC	AAC49819	standard; DNA; 504 BP.	
AC	AAC49819;		
DT	18-OCT-2000	(first entry)	
XX		Arabidopsis thaliana DNA fragment SEQ ID NO: 62553.	
DE		Hybridisation assay; genetic mapping; gene expression control;	
KW		protein identification; signal transduction pathway; metabolic pathway;	
KW		promoter; termination sequence; ss.	
XX		Arabidopsis thaliana.	
OS			
XX			
XX	EP1033405-A2.		
XX			
PD	06-SEP-2000.		
XX			
XX	25-FEB-2000; 2000EP-00301439.		
PR	25-FEB-1999;	99US-0121825P.	
PR	03-MAR-1999;	99US-0123180P.	
PR	09-MAR-1999;	99US-0123548P.	
PR	23-MAR-1999;	99US-0125788P.	
PR	25-MAR-1999;	99US-0126264P.	
PR	29-MAR-1999;	99US-0126785P.	
PR	01-APR-1999;	99US-0127462P.	
PR	08-APR-1999;	99US-0128234P.	
PR	16-APR-1999;	99US-0128714P.	
PR	19-APR-1999;	99US-0129845P.	
PR	21-APR-1999;	99US-0130077P.	
PR	23-APR-1999;	99US-0130449P.	
PR	28-APR-1999;	99US-0130510P.	
PR	28-APR-1999;	99US-0131449P.	
PR	30-APR-1999;	99US-0132048P.	
PR	30-APR-1999;	99US-0132407P.	
PR	04-MAY-1999;	99US-0132484P.	
PR	05-MAY-1999;	99US-0132485P.	
PR	06-MAY-1999;	99US-0132486P.	
PR	07-MAY-1999;	99US-0132487P.	
PR	07-MAY-1999;	99US-0132863P.	
PR	11-MAY-1999;	99US-0134256P.	
PR	14-MAY-1999;	99US-0134218P.	
PR	14-MAY-1999;	99US-0134219P.	
PR	14-MAY-1999;	99US-0134221P.	
PR	14-MAY-1999;	99US-0134370P.	
PR	18-MAY-1999;	99US-0134768P.	
PR	19-MAY-1999;	99US-0134941P.	
PR	20-MAY-1999;	99US-0135124P.	
PR	21-MAY-1999;	99US-0135353P.	
PR	24-MAY-1999;	99US-0135629P.	
PR	25-MAY-1999;	99US-0136021P.	
PR	27-MAY-1999;	99US-0136392P.	
PR	28-MAY-1999;	99US-0136782P.	
PR	01-JUN-1999;	99US-0137222P.	
PR	03-JUN-1999;	99US-0137528P.	
PR	04-JUN-1999;	99US-0137502P.	
PR	07-JUN-1999;	99US-0137724P.	
PR	08-JUN-1999;	99US-0138094P.	
PR	10-JUN-1999;	99US-0138540P.	
PR	10-JUN-1999;	99US-0138847P.	
PR	14-JUN-1999;	99US-0139119P.	
PR	16-JUN-1999;	99US-0139452P.	
PR	16-JUN-1999;	99US-0139453P.	
PR	17-JUN-1999;	99US-0139492P.	
PR	18-JUN-1999;	99US-0139454P.	
PR	18-JUN-1999;	99US-0139455P.	
PR	18-JUN-1999;	99US-0139456P.	
PR	18-JUN-1999;	99US-0139457P.	
PR	18-JUN-1999;	99US-0139458P.	
PR	18-JUN-1999;	99US-0139459P.	
PR	18-JUN-1999;	99US-0139460P.	
PR	18-JUN-1999;	99US-0139461P.	
PR	18-JUN-1999;	99US-0139462P.	
PR	18-JUN-1999;	99US-0139463P.	
PR	18-JUN-1999;	99US-0139750P.	
PR	18-JUN-1999;	99US-0139763P.	
PR	21-JUN-1999;	99US-0139817P.	
PR	22-JUN-1999;	99US-0139899P.	
PR	23-JUN-1999;	99US-0140353P.	
PR			

PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161992P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.	Alignment Scores:		
PR	05-AUG-1999;	99US-0147192P.	Pred. No.:	1.69e-26	Length: 504
PR	05-AUG-1999;	99US-0147260P.	Score:	277.50	Matches: 59
PR	06-AUG-1999;	99US-0147303P.	Percent Similarity:	69.8%	Conservative: 22
PR	06-AUG-1999;	99US-0147416P.	Best Local Similarity:	50.9%	Mismatches: 25
PR	09-AUG-1999;	99US-0147493P.	Query Match:	42.3%	Indels: 10
PR	09-AUG-1999;	99US-0147935P.	DB:	3	Gaps: 2
PR	10-AUG-1999;	99US-0148171P.	US-10-697-787-2 (1-126) x AAC49819 (1-504)		
PR	11-AUG-1999;	99US-0148119P.	QY	1	MetAlaSerAsnSerArgSerSerIleSerProTTPThrPheSerGlnAenLysMetPhe 20
PR	12-AUG-1999;	99US-0148341P.	Db	103	ATGGCTTCCAACCTCAATGAGCTCTAGCGCTTCTTGACACACGTAAAGGAGAACAAATATTT 162
PR	13-AUG-1999;	99US-0148565P.	QY	21	GluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTTPHisAenValAlaLys 40
PR	13-AUG-1999;	99US-0148684P.	Db	163	GAAAGGGCGTTGGCTACATATATGACAGGACACTCTGACCGTTGGCATACGTTGCAAGA 222
PR	16-AUG-1999;	99US-0149368P.	QY	41	AlaValGlyGlyLysThrValGluGluValLysArgHisTyrAsp-IleLeuValGluAs 60
PR	17-AUG-1999;	99US-0149175P.	Db	223	GCGTTGGCGCAAAATCAGCTGAAGAAGTAAGCGACACTACGAGCTCTCTATTAGGGA 282
PR	18-AUG-1999;	99US-0149426P.	QY	60	pLeuIleAsnIleGluThrGlyArgValProLeuProAenTyrLysThr-PheGluSerA 80
PR	20-AUG-1999;	99US-0149722P.	Db	283	TGTCAATGACATTGAGTCAGGCGCTTATCCACATCCCAATTACCGTTCAAATGGAAACAA 342
PR	20-AUG-1999;	99US-0149723P.	QY	80	snSerArgSerIle-AsnAspPheAspThrArgTyrIleThrLysTyrLeuTyrMetMet 99
PR	20-AUG-1999;	99US-0149929P.	Db	343	CCACTGAAGCATTAAGGAACCTCCAA---AAGGCTCTACTGAATGG----- 387
PR	23-AUG-1999;	99US-0149902P.	QY	100	LeuSerIleTyrPheAspAenHisSerSerAspPheGluLys 113
PR	23-AUG-1999;	99US-0149930P.	Db	388	-----ACATGGTTCCAAAACCATTCACAAAGCACATCTAAG 423
PR	25-AUG-1999;	99US-0150566P.	RESULT 15		
PR	26-AUG-1999;	99US-0150884P.	ADV43786/C	ADJ43786 standard; cDNA; 466 BP.	
PR	27-AUG-1999;	99US-0151065P.	ID	ADJ43786 standard; cDNA; 466 BP.	
PR	27-AUG-1999;	99US-0151066P.	XX	ADJ43786;	
PR	27-AUG-1999;	99US-0151080P.	AC	06-MAY-2004 (first entry)	
PR	30-AUG-1999;	99US-0151303P.	XX	Plant cDNA #4786.	
PR	31-AUG-1999;	99US-0151438P.	DE	Plant; gene; ss; transcription; plant genome augmentation; cereal;	
PR	01-SEP-1999;	99US-0151930P.	XX	soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;	
PR	07-SEP-1999;	99US-0152363P.	KW	maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;	
PR	10-SEP-1999;	99US-0153070P.	KW	stresses tolerance; salt tolerance; cold tolerance; drought tolerance;	
PR	13-SEP-1999;	99US-0153758P.	KW	plant nutrition; apical dominance; dwarfism; early flowering; antiviral;	
PR	15-SEP-1999;	99US-0154018P.	XX	antifungal.	
PR	16-SEP-1999;	99US-0154019P.	OS	Eukaryota.	
PR	20-SEP-1999;	99US-0154779P.	XX	US2004016025-A1.	
PR	22-SEP-1999;	99US-0155139P.	XX	22-JAN-2004.	
PR	23-SEP-1999;	99US-0155486P.	XX	26-SEP-2002; 2002US-00260238.	
PR	24-SEP-1999;	99US-0155659P.	XX	26-SEP-2001; 2001US-0325277P.	
PR	28-SEP-1999;	99US-0156458P.	PR	26-SEP-2001; 2001US-0325448P.	
PR	29-SEP-1999;	99US-0156596P.	PR	04-APR-2002; 2002US-0370620P.	
PR	04-OCT-1999;	99US-0157117P.	XX	(BUDW/) BUDWORTH P.	
PR	05-OCT-1999;	99US-0157533P.	PA	(MOUG/) MOUGHAMER T.	
PR	06-OCT-1999;	99US-0157865P.			
PR	07-OCT-1999;	99US-0158029P.			
PR	08-OCT-1999;	99US-0158232P.			
PR	12-OCT-1999;	99US-0158369P.			
PR	13-OCT-1999;	99US-0159293P.			
PR	13-OCT-1999;	99US-0159294P.			
PR	13-OCT-1999;	99US-0159295P.			
PR	14-OCT-1999;	99US-0159329P.			
PR	14-OCT-1999;	99US-0159330P.			
PR	14-OCT-1999;	99US-0159331P.			
PR	14-OCT-1999;	99US-0159637P.			
PR	14-OCT-1999;	99US-0159638P.			
PR	18-OCT-1999;	99US-0159584P.			
PR	21-OCT-1999;	99US-0160741P.			
PR	21-OCT-1999;	99US-0160767P.			
PR	21-OCT-1999;	99US-0160768P.			
PR	21-OCT-1999;	99US-0160770P.			
PR	21-OCT-1999;	99US-0160814P.			
PR	21-OCT-1999;	99US-0160815P.			
PR	22-OCT-1999;	99US-0160980P.			
PR	22-OCT-1999;	99US-0160981P.			
PR	22-OCT-1999;	99US-0160989P.			
PR	25-OCT-1999;	99US-0161404P.			
PR	25-OCT-1999;	99US-0161405P.			
PR	25-OCT-1999;	99US-0161406P.			

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, **using frame_plus_p2n model**

Run on: March 10, 2006, 19:17:45 ; Search time 3944 Seconds

(without alignments)
1815.993 Million cell updates/sec

Title: US-10-697-787-2

Perfect score: 656

Sequence: 1 MASNRSSISPTWFSQNMKFM.....HSSDFEKFQKVLVSIVISLV 126

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh
-Q=/abs/ABSNWEB_spool/US10697787/runat_09032006_150627_3319/app_query.fasta_1
-B=GenEmbl -QMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs08
-USER=US10697787 @CGN 1.1.4919 @runat_09032006_150627_3319 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.in.*
3: gb.env.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pr.*
9: gb.ro.*
10: gb.sts.*
11: gb.sv.*
12: gb.un.*
13: gb.vi.*
14: gb.hcg.*
15: gb.pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	656	100.0	381	15 AY519525	Arabidopsis
2	656	100.0	68041	15 AC025814	Arabidopsis
3	473	72.1	294	6 CS138004	Sequence

4	344.5	52.5	334	15	BT005657	Arabidops
5	344.5	52.5	580	15	AK119034	Arabidops
6	343	52.3	303	15	AY519524	Arabidops
C 7	343	52.3	120977	15	AC025808	Genomic s
C 8	296	45.1	54433	15	AC007119	Arabidops
C 9	293	44.7	108844	15	AC157502	Medicago
10	289	44.1	246	15	BT011255	Arabidops
11	289	44.1	506	15	BT010770	Arabidops
12	286	43.6	306	15	AY519526	Arabidops
13	286	43.6	337	15	BT008554	Arabidops
14	286	43.6	570	15	BT008698	Arabidops
15	286	43.6	578	6	CS137904	Sequence
16	281	42.8	100108	14	AP008111	Lotus cor
C 17	276.5	42.1	89035	15	ATT22F8	Arabidops
C 18	276.5	42.1	198987	15	ATCHRIV90	Arabidops
19	276	42.1	3213	15	AY54971	Antirrhin
C 20	276	42.1	198354	15	ATAP22	Arabidops
21	276	42.1	198750	15	ATAP22	Arabidops
22	275.5	42.0	294	6	CS137908	Sequence
23	275.5	42.0	294	15	AY519527	Arabidops
24	274.5	41.8	3605	15	AJ583670	Lycopersi
25	263	40.1	612	15	LES277944	Lycopersi
C 26	261.5	39.9	97268	15	AP004546	Lotus cor
C 27	261.5	39.9	102550	14	AP008039	Lotus cor
C 28	261.5	39.9	108878	14	AP007808	Lotus cor
C 29	260	39.6	115071	14	AC135566	Medicago
C 30	254.5	38.8	166907	14	AC148974	Zea mays
31	252	38.4	512	15	AK112054	Oryza sat
32	252	38.4	34070	15	AP007251	Oryza sat
33	252	38.4	110000	15	AP008207	Continuation (255
C 34	252	38.4	110220	14	AP003688	Oryza sat
C 35	252	38.4	162297	15	AP004317	Oryza sat
C 36	251	38.3	91322	14	AP007301	Lotus cor
C 37	246.5	37.6	99587	15	AC007858	Oryza sat
C 38	246.5	37.6	110000	15	AP008211	Continuation (287
C 39	246.5	37.6	156643	15	AC120988	Oryza sat
C 40	245.5	37.4	92861	15	AC006439	Arabidops
41	245	37.3	513	6	CS137972	Sequence
42	244	37.2	72767	14	AP007919	Lotus cor
43	241	36.7	108564	14	AC162787	Medicago
44	218.5	33.3	110000	15	AP008218	Continuation (206
C 45	218.5	33.3	143618	15	CNS08C7N	Oryza sat

ALIGNMENTS

RESULT 1	AY519525	Arabidopsis thaliana MYB transcription factor (Atig575250) mRNA,	381 bp	linear	PLN 07-FEB-2004
LOCUS	AY519525	complete cds.			
DEFINITION	Arabidopsis thaliana				
ACCESSION	AY519525				
VERSION	AY519525.1	GI:41618977			
KEYWORDS	Arabidopsis thaliana (thale cress)				
SOURCE	Arabidopsis thaliana				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi				
REFERENCE	1 (bases 1 to 381)				
AUTHORS	Qu,L. and Gu,H.				
TITLE	The MYB Transcription Factor Family in Arabidopsis: A Genome-Wide Cloning and Expression Pattern Analysis				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 381)				
AUTHORS	Qu,L. and Gu,H.				
TITLE	Direct Submission				
JOURNAL	Submitted (07-JAN-2004) Life Sciences, National Laboratory of Protein Engineering and Plant Genetic Engineering, Peking University, Beijing 100871, China				
FEATURES	Location/Qualifiers				
source	1..381				
	/organism="Arabidopsis thaliana"				

/mol_type="mRNA"
/db_xref="taxon:3702"
1. .381
/locus_tag="At1g75250"
1. .381
/locus_tag="At1g75250"
/codon_start=1
/product="MYB transcription factor"
/protein_id="AA090995.1"
/db_xref="GI:41618978"
/translations="MASNSRSPSSIPWTFQKNKFERALAVYDKTDPRHNVAKAVGG
KTVBEVKRHYDIIVELDLINIEORVPLPNKYKTFESRSRINDPDTYIKLYLMLLSI
YFDNHSSDFEKFQKLVSYISLV"

ORIGIN

Alignment Scores:
Pred. No.: 5, 92e-63 Length: 381
Score: 656.00 Matches: 126
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 15 Gaps: 0

US-10-697-787-2 (1-126) x AY519525 (1-381)

Qy 1 MetAlaSerAenSerArgSerSeriLeSerProTrpThrPheSerGlnAenLysMetPhe 20
Db 1 ATGGCGTCAAACTCAAGAGTTCAATCTCACCATGGAGCTTTAGTCAAAACAAGATGTTTC 60
Qy 21 GluArgAlaLeuAlaValTyAspLysAspThrProAspArgTrpHisAenValAlaLys 40
Db 61 GAGAGGCGCTTGGCAGTTTACGACAGGACACACCCGACCGATGGCAATGTGGCAAAA 120
Qy 41 AlaValGlyGlyLysThrValGluGluValLysArgHisTrpAspLysLeuValGluAsp 60
Db 121 GCTGTGCGAGGAGAACTGTAGAGAGTGAAGCGCCACTATGACATCTCGTGGAGAT 180
Qy 61 LeuLeuAenLeuGluThrGlyArgValProLeuProAenTyriLysThrPheGluSerAen 80
Db 181 CTCATCAACATCGAGACTGGTGTGTCCTTTGCCAATTAACAAGACTTCGAATCTTAAC 240
Qy 81 SerArgSerileAenAspPheAspThrArgTyriLeuThrLysTyriLeuTyriMetLeu 100
Db 241 TCAAGAGCATCAATGACTTTGACACAGGATGATATATACTAAATCTATATATGATGCTC 300
Qy 101 SerileTyriPheAspAenHisSerSerAspPheGluLysPheSerGlnLysValLeuVal 120
Db 301 TCGATATATTTTGATATCATCTAGTGATTTTGAGAAATCTCTCAAAAAGTTCCTGTA 360
Qy 121 SerTyriLeuSerLeuVal 126
Db 361 AGTTATATTTCTTTGGTT 378

RESULT 2
AC025814
LOCUS Arabidopsis thaliana chromosome 1 BAC F22H5 genomic sequence, complete sequence.
DEFINITION AC025814.7 GI:12331602
ACCESSION
VERSION
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 68041)
AUTHORS Lin, X., Kaul, S., Town, C.D., Benito, M.-I., Creasy, T.H., Haas, B.J., Wu, D., Maiti, R., Ronning, C.M., Koo, H., Fujii, C.Y., Utterback, T.R., Barnstead, M.E., Bowman, C.L., White, O., Nierman, W.C. and Fraser, C.M.
TITLE Arabidopsis thaliana chromosome 1 BAC F22H5 genomic sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 68041)

AUTHORS
TITLE Direct Submission
JOURNAL
REFERENCE 3 (bases 1 to 68041)
AUTHORS Town, C.D. and Kaul, S.
TITLE Direct Submission
JOURNAL
REFERENCE 4 (bases 1 to 68041)
AUTHORS Town, C.D. and Kaul, S.
TITLE Direct Submission
JOURNAL
REFERENCE 5 (bases 1 to 68041)
AUTHORS Town, C.D. and Kaul, S.
TITLE Direct Submission
JOURNAL
REFERENCE 6 (bases 1 to 68041)
AUTHORS Town, C.D. PhD.
TITLE Direct Submission
JOURNAL
COMMENT On Jan 22, 2001 this sequence version replaced gi:12280794.
Address all correspondence to: at@tigr.org

Lin, X. and Kaul, S.

Direct Submission
Submitted (15-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, xlin@tigr.org
3 (bases 1 to 68041)
Town, C.D. and Kaul, S.

Direct Submission
Submitted (12-SEP-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
4 (bases 1 to 68041)
Town, C.D. and Kaul, S.

Direct Submission
Submitted (05-OCT-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
5 (bases 1 to 68041)
Town, C.D. and Kaul, S.

Direct Submission
Submitted (12-OCT-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
6 (bases 1 to 68041)
Town, C.D. PhD.

Direct Submission
Submitted (22-JAN-2001)
On Jan 22, 2001 this sequence version replaced gi:12280794.
Address all correspondence to: at@tigr.org

BAC clone F22H5 is from Arabidopsis thaliana chromosome 1
The orientation of the sequence is from SP6 to T7 end of the BAC
clone.

Genes were identified by a combination of several methods: Gene
prediction programs including Gensecan+ (Chris Burge,
http://CCR-081.mit.edu/GENSECAN.html), GeneMarkHM (Mark Borodovsky,
http://genemark.biology.gatech.edu/GeneMark/), GlimmerA (a variant
of Glimmer, see Mihaela Pertea,
http://www.tigr.org/softlab/glimmerm.htm/glimmerm.html, and
GeneSplicer (Mihaela Pertea and Steven Salzberg, contact
mpertea@tigr.org), searches of the complete sequence against a
peptide database and the plant EST database at TIGR
(http://www.tigr.org/tdb/tgi.shtml). Annotated genes are named to
indicate the level of evidence for their annotation. Genes with
similarity to other proteins are named after the database hits.
Genes without significant peptide similarity but with EST
similarity are named as unknown proteins. Genes without protein
or EST similarity, that are predicted by more than two gene
prediction programs over most of their length are annotated as
hypothetical proteins. Genes encoding tRNAs are predicted by
tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/).
Simple repeats are identified by RepeatMasker (Arian Smit,
http://ftp.genome.washington.edu/RM/RepeatMasker.html).

FEATURES
source

1. .68041
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="F22H5"
/ecotype="Columbia"
complement(418..5404)
/genes="F22H5.15"
complement(join(<418..594,684..773,873..938,1034..1095,
1179..1227,1319..4820,5040..5404))
/gene="F22H5.15"
complement(join(<418..594,684..773,873..938,1034..1095,
1179..1227,1319..4820,5040..5404))
/genes="F22H5.15"
/codon_start=1
/product="unknown protein, 3' partial; 5120-134"
/protein_id="AA012690.1"
/db_xref="GI:10092277"
/translocation="MEYKSPATTTTSRKISNNRSHLSFSANAVDGVSSPVSNSKS
PLVDYGEI FRGGPSPSSIPFLDVLNVLGVKVDVRSSKLDYSSVFGGACDAFVT
PKVEI IKSCKTSINEDKRRKGGNSDVLPLCNKGKSPKPMVRMKHSDISYHQTVP
RNENGATHLTQVPATPGPIPTQVDVNTSLHLKIESKSTPIPAVEKKLPCNREGEVK

```

AGRKQSGKTEVDFENIFARDGCGSTRDDSTCKTSGNGEYRDVKPPSPFQCTLNGEHCAS
ERLGLNSGSPSERVEDADSPSPYPYFDAETKDENSVAESAALKKAIEEAQIRWNI
AKOMMEKSGFRSCAKLSCDDSKIENKGTNVKGIITEESRDNSQILIEEAQIRWNI
SFSNEGDOHAKRKLWGPGLLKSTSDHKPELEEQDIITLEEQARRGRKHNEP
GLIFKSYMNSKQEPENLAPAKPEPTDKOEVQITENPIYTFQGLSKLKCVCVEAFTG
SKVSKQKQFTKEKENSITVQVQDEESQEMLAGIPVETYLREVESTPOOTESKS
ENWIEKESPTICAFTESSONMEKETGMQVKSACKFEDSGVKDFQFENGDDPCNVLD
QGEKEIIVSEPOEMLVPDSDSKTYVREBEETPTPSLNTQSDSDSGVAMVSNRWVISE
PGNIDEBQAVHKVPRRRVWKTSDVNNMIKAPKGNRPMWOLESAENETNAMSHEE
GVRIHASFEISTSQASDSDGLQENWTVLKQFMQFQATDKGEDETTCYLVESERG
HLDIHQAKQKEQVEVEVTRVNYEAHATRENDESAQETTCRKEDGKVEVQGTSL
VRELLGEOLMASLEEDVQEAEEAGWVQGLSELMNEINHADSHAEMLLEYDRSTDS
NNSRFRFQTOQAEBETWIDGSDITDTSRSPFMQGDYSIIEVIEGHRSDQFPEKA
SAVNTNEHIEBIDSISQSGWSVVEDDSDSQDQGAOESKHDELBEATKEESDEM
TSLGVERNGDKELHOPCEQWNETYRSNVAAESSRPNNGEEIIGAATNGMKE
GREGESCRSMEEGSDATSDIGAATDGMKNEGEESCRSMEEGSDATSDISQVKA
ETVEHLLKIDITREKERERKQVWVERAIREARAPADARAGKTAMEKAKAVA
HREVPKSGESVEVNDKLSSAERKASQAKURAAVRAITETVRERAMEKALSGK
SAASQAKYGGSKFSFGSSGERRSSSTENKSGPSSNQTGPICRQKARSRHQ
RTSDRAALAAEKLRDLKTOKEQTERNRLEALDADVKRWSGKNNLRALISTLOY
ILGAESGWKPIPLTDLVSSASVRKAYRKATLVHPDKLQORGASTQOKYICEKVFLL
K"
complement(9354..10899)
/gene="F22H5.16"
/note="similar to GI:1708420 from [Arabidopsis thaliana]
(J. Biol. Chem. 270 (44), 26224-26231 (1995))"
complement(join(<9354..9638,9750..9953,10043..10328,
10461..10595,10671..>10899))
/gene="F22H5.16"
complement(join(9481..9638,9750..9953,10043..10328,
10461..10595,10671..10856))
/gene="F22H5.16"
/codon_start=1
/product="NADPH oxidoreductase, putative; 10572-9197"
/protein_id="AAG12695.1"
/db_xref="GI:10092282"
/translation="MTTEKSKILVIGGTGYMGFEIVGESAKAGNPTFALVREASISDP
VKSRTIQSKDLGVTLHGLDNHESLVKAIKQDVVISTGHKQIFQDTKIIAISAIKE
AGNVKRFILPABFGIVERTSAVEPAKSLFAGKVQIRRAIEAEGIPYTVVNSKSAFY
LTLQLQSGSLISHPRDKAIIFGDKVPPRDKVTILGDNKAVINKEDVAAYNIKA
VDDLTLNKTLYISPRNLLSNMENVTLWEKIGKSLKTHISEQILKSIQVPIDVF
KSINHAVVKGDTSTFTIEPMPFGEASVLYPDVKYTSIDVLSQPT"
complement(11122..12573)
/gene="F22H5.18"
/note="similar to GI:1708420 from [Arabidopsis thaliana]
(J. Biol. Chem. 270 (44), 26224-26231 (1995))"
complement(join(<11122..11422,11516..11734,11791..12034,
12117..12251,12333..>12573))
/gene="F22H5.18"
complement(join(11235..11422,11516..11734,11791..12034,
12117..12251,12333..12518))
/gene="F22H5.18"
/codon_start=1
/product="NADPH oxidoreductase, putative; 12234-10951"
/protein_id="AAG12680.1"
/db_xref="GI:10092267"
/translation="WASBKSILVIGGTGHIGKLIIEASVKAGHSTLALVREASISDP
NKGKTVQKDFGVTLHGLDNHESLVKAIKQADVISTVSGMQILDQTKIIAISAIKE
AGNVKRFILPBFSGMDVKSASVEPAKSAFGRKLQTRDIEAEGIPYTVLVNIFAGY
LPTVLQLEPFGTSPRDKVIFGDNKVVEYFIYAVINKEDIAAYTIKAVDDPRTLK
LTINPNPNTLSMNEIVTLWEKIGKSLKTHISEQILKSIQVPSVFNLLSINHA
VFKVGDQTNFTIEPFGPEASELYPDIKYTSIDVLSYFALGTSINT"
complement(12872..14426)
/gene="F22H5.17"
/note="similar to GI:1708420 from [Arabidopsis thaliana]
(J. Biol. Chem. 270 (44), 26224-26231 (1995))"
complement(join(<12872..13216,13298..13513,13578..13821,
13974..14108,14193..>14426))
/gene="F22H5.17"
complement(join(13053..13216,13298..13513,13578..13821,
13974..14108,14193..14378))
/gene="F22H5.17"
/codon_start=1

```

```

/product="NADPH oxidoreductase, putative; 14094-12769"
/protein_id="AAG12677.1"
/db_xref="GI:10092264"
/translation="WATEKSKILVIGGTGYIGKLIIEASAKAGHSTFALVREATISDP
VKGKTVQKDFGVTLHGLDNHESLVKAIKQADVISTVSGMQILDQTKIIAISAIKE
AGNVKRFILPBFSGMDVKSASVEPAKSAFAGKIQIRRTIEAEGIPYTVAVTGFSGGY
LPTVLQLEPFGTSPRDKVIFGDNKANEYFAVINKEDIAAYTIKAVDDPRTLKNI
LVTKPQNTLSMNEIVTLWEKIGKSLKTHISEQILKSIQVPSVFNLLSINHA
VFKVGDQTNFTIEPFGPEASELYPDVKYTSVDYLSYFA"
complement(15081..16153)
/gene="F22H5.1"
/note="similar to GI:6939839 from [Oryza sativa]
(join(15081..15228,15380..15470,15661..>16153))
/gene="F22H5.1"
complement(join(15171..15228,15380..15470,15661..16153))
/gene="F22H5.1"
/codon_start=1
/product="GSH-dependent dehydroascorbate reductase 1,
putative; 14887-15869"
/protein_id="AAG12679.1"
/db_xref="GI:10092266"
/translation="WALDICKVAVGADVLGDCPFSORVLLTLEKKLPYKTHLINV
SDKPQWFLDISPEGKVPVYKLDGKWADSDVIVGLLEEKYPPEPKTPPEFASVSGKI
FGAFVTLKSKDANDSGSEKALVDELEALENLKTHSGFPVAGEKITAVDLSLAPKLYH
LEVALGHYKNWSPESLTSVRNKAALFRESFENTAKKEIIVWAGWESKVNA"
complement(16410..18554)
/gene="F22H5.2"
complement(join(<16410..16558,16703..16830,17050..>18554))
/gene="F22H5.2"
complement(join(16410..16558,16703..16830,17050..18554))
/gene="F22H5.2"
/codon_start=1
/product="unknown protein; 18270-16126"
/protein_id="AAG12682.1"
/db_xref="GI:10092269"
/translation="MPPFRFRIPFPSSNSRISSTSSPSPPTPPSPRPPRPAGIA"

```

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:
3,156-60	68041	126	
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DBs:	15	Gaps:	0

US-10-697-787-2 (1-126) x AC025814 (1-68041)

Qy	1	MetAlaSerAsnSerArgSerSerIleSerProTprThrPheSerGlnAsnLysMetPhe	20
Db	20955	ATGGCGTCAAACTCAAGAAGTTCAATCTCACCATGGAGCGTTAGTCAAAACAAGATGTC	21014

Qy	21	GluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTprPheAsnValalalys	40
Db	21015	GAGAGGCGCTTGGCAGTTTACGACAAAGGACACACCGACCGATGGCACAATGTGCAAAA	21074

Qy	41	AlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleleuValGluasp	60
Db	21075	GCTGTGGAGGAGAAACTGTAGAAGAGTGAAGCCCACTATGACATCTCTCGCAGGAT	21134

Qy	61	LeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGluSerAsn	80
Db	21135	CTCATCAACATCGAGACGTGTCGTCCTTTGGCCATTACACACCTTCGAATCAAC	21194

Qy	81	SerArgSerIleAsnAspPheAspThrArgTyrIleThrLysTyrLeuTyrMetMetLeu	100
Db	21195	TCAAGAAGCATCAATGACTTTGACACAAAGGTATATACTAATAATCTATATATGATGCTC	21254

Qy	101	SerIleTyrPheAspAsnHisSerSerAspPheGluLysPheSerGlnLysValleuVal	120
Db	21255	TCGATATATTTTGATAATCATCTAGTGATTTTGAGAAATTTCTCTCAAAAAGTCTCTGTA	21314

Qy	121	SerTyrIleSerLeuVal	126
Db	21314		


```
|||||...:|||||...|||||...|||||...|||||...|||||...
Db 1 ATGGCTCTAGTTCTTCTAGAGCTCGAGCTCTTCTTGGACGCTCTAAGCAAAACAAGATGTTTC 60
QY 21 GUAAGAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisValAlaLys 40
Db 61 GAGAGGGCTTTACCGTTTACGATAAAGACACTCCCGACCGTTGGCAAAACGCGCTAAA 120
QY 41 AlaValGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAsp 60
Db 121 GCAGTTGGAAGTAAATCTGCAGAGGAAGTTAAACGTCCTACGACATCCTCGTTGAAGAT 180
QY 61 LeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGlu----- 78
Db 181 CTCATGAACATCGAACCAAGACTTAGTACCTTTTCCTTAAATACAAAACCGTCGATGTGGA 240
QY 79 SerAsnSerArgSerIleAsnAspPheAspThrArgTyrIleThrLysTyr----- 95
Db 241 AGTAAATCTAGAGGCATCGATGTTTGTAGGTAAATGAAGAATATGAGAAATCCAG 300
QY 96 -----LeuTyrMetMet 99
Db 301 TGAAGTTGTTACATGATG 318

AKI19034 580 bp mRNA linear PLN 14-FEB-2004
Arabidopsis thaliana At1g19510 mRNA for putative myb-related
protein, complete cds, clone: RAF121-36-F10.
AKI19034
AKI19034.1 GI:26453067
FLI CDNA; CAP trapper.
Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Arabidopsi.
1
Seki,M., Iida,K., Satou,M., Sakurai,T., Akiyama,K., Ishida,J.,
Nakajima,M., Enju,A., Kamiya,A., Narusaka,M., Carninci,P.,
Kawai,J., Hayashizaki,Y. and Shinozaki,K.
Arabidopsis thaliana full-length cDNA
Published Only in Database (2002)
2 (bases 1 to 580)
Seki,M., Iida,K., Satou,M., Sakurai,T., Akiyama,K., Ishida,J.,
Nakajima,M., Enju,A., Kamiya,A., Narusaka,M., Carninci,P.,
Kawai,J., Hayashizaki,Y. and Shinozaki,K.
Direct Submission
Submitted (25-NOV-2002) Motoaki Seki, RIKEN Genomic Sciences
Center; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa
230-0045, Japan (E-mail:mseki@gsc.riken.go.jp,
URL:http://pfweb.gsc.riken.go.jp, Tel:81-45-503-9625,
Fax:81-45-503-9586)
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al. (1998) Plant J. 15:707-720;
Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda FLC-1-E vector (Carninci et
al. (2001) Genomics 77:79-90) digested with BamHI and SalI.
This clone is in a modified pBluescript vector.
Please visit our web site (http://pfweb.gsc.riken.go.jp/) for
further details.
Location/Qualifiers
1..580
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAF121-36-F10"
/ecotype="Columbia"
/note="common name: thale cress"
1..580
/gene="At1g19510/F18014_31"
88..390
/gene="At1g19510/F18014_31"

RESULT 5
AKI19034
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
1..580
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAF121-36-F10"
/ecotype="Columbia"
/note="common name: thale cress"
1..580
/gene="At1g19510/F18014_31"
88..390
/gene="At1g19510/F18014_31"
```

```
/codon_start=1
/produced="putative myb-related protein"
/protein_id="BAC43610.1"
/db_xref="GI:26453068"
/translation="MASSMSSSSSWTSKQKMFERALAVYDKDTPDRMKNVAKVGS
KSABEVKGRHYDILVEDLMNIEQDLVPLPKYKTVGVSGSRGIDDFDLRLMKNRIQ"

ORIGIN
Alignment Scores:
Pred. No.: 1..58e-28 Length: 580
Score: 344.50 Matches: 72
Percent Similarity: 73.6% Conservative: 6
Best Local Similarity: 67.9% Mismatches: 21
Query Match: 52.5% Indels: 7
DB: 15 Gaps: 2

US-10-697-787-2 (1-126) x AKI19034 (1-580)
QY 1 MetaIaSerAsnSerArgSerIleSerProTyrThrPheSerGlnAsnLysMetPhe 20
Db 88 ATGGCTCTAGTTCTTCTAGAGCTCGAGCTCTTCTTGGACGCTCTAAGCAAAACAAGATGTTTC 147
QY 21 GUAAGAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisValAlaLys 40
Db 148 GAGAGGGCTTTAGCGTTTACGATAAAGACACTCCCGACCGTTGGCAAAACGCGCTAAA 207
QY 41 AlaValGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAsp 60
Db 208 GCAGTTGGAAGTAAATCTGCAGAGGAAGTTAAACGTCCTACGACATCCTCGTTGAAGAT 267
QY 61 LeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGlu----- 78
Db 268 CTCATGAACATCGAACCAAGACTTAGTACCTTTGCCCTTAAATACAAAACCGTCGATGTGGA 327
QY 79 SerAsnSerArgSerIleAsnAspPheAspThrArgTyrIleThrLysTyr----- 95
Db 328 AGTAAATCTAGAGGCATCGATGTTTGTAGGTAAATGAAGAATATGAGAAATCCAG 387
QY 96 -----LeuTyrMetMet 99
Db 388 TGAAGTTGTTACATGATG 405

RESULT 6
AKI19524
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..303
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
1..303
/locus_tag="At1g19510"

AY519524 303 bp mRNA linear PLN 07-FEB-2004
Arabidopsis thaliana MYB transcription factor (At1g19510) mRNA,
complete cds.
AY519524
AY519524.1 GI:41618973
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Arabidopsi.
1 (bases 1 to 303)
Qu,L. and Gu,H.
The MYB Transcription Factor Family in Arabidopsis: A Genome-Wide
Cloning and Expression Pattern Analysis
Unpublished
2 (bases 1 to 303)
Qu,L. and Gu,H.
Direct Submission
Submitted (07-JAN-2004) Life Sciences, National Laboratory of
Protein Engineering and Plant Genetic Engineering, Peking
University, Beijing 100871, China
Location/Qualifiers
1..303
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
1..303
/locus_tag="At1g19510"
```

CDS	1. .303 /locus tag="At1g19510" /codon_start=1 /product="MYB transcription factor" /protein_id="AA509994.1" /db_xref="GI:41618974" /translation="MASSNSSSSSWTSKMKFERALAVYKDTDPDRQNVAKVAGS KSAEVRKRYDILVEDLMNIEODLPKPKYKTVGVGSKRGIDDFDLRLMKNNRIQ"	JOURNAL	Submitted (22-APR-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
ORIGIN		REFERENCE AUTHORS	4 (bases 1 to 120977) Cheuk,R., Shinn,P., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bei,B., Chin,C., Chiou,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,S., Mukharasky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vayseberg,M., Yu,G., Davis,R., Federspiel,N., Theologis,A. and Ecker,J.
Alignment Scores:		TITLE	Direct Submission
Pred. No.:	1.05e-28	JOURNAL	Submitted (28-JUN-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA
Score:	343.00	REFERENCE	5 (bases 1 to 120977)
Percent Similarity:	80.4%	AUTHORS	Chao,Q., Brooks,S., Buehler,E., Johnson-Hopson,C., Khan,S., Kim,C., Shinn,P., Altafi,H., Bei,B., Chin,C., Chiou,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vayseberg,M., Yu,G., Davis,R., Federspiel,N., Theologis,A. and Ecker,J.
Best Local Similarity:	75.0%		Direct Submission
Query Match:	52.3%		Submitted (11-OCT-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA
DB:	15		On Apr 22, 2000 this sequence version replaced gi:7543634.
US-10-697-787-2 (1-126) x AY519524 (1-303)		TITLE	Location/Qualifiers
Qy 1 MetAlaSerAsnSerArgSerIleSerProTrpThrPheSerGlnAsnLysMetPhe 20		JOURNAL	1. .120977 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /db_xref="taxon:3702" /chromosome="1" /clone="F18014" 90. .2159 /notes="hypothetical protein"
Db 1 ATGGCTCTAGTTCTATGAGCTCGAGCTCTTCTTGGACGCTTAAGCAAAACAAGATGTTTC 60		source	/codon_start=1 /evidence="not experimental" /product="F18014.1" /protein_id="AAF79419.1" /db_xref="GI:8778411"
Qy 21 GluArgAlaLeuAlaValTyAspLysAspThrProAspArgTrpHisAsnValAlaLys 40		CDS	complement(4647..5702) /notes="putative glycosyltransferase-like protein emb CAB42905.1; similar to EST gb A1998490.1"
Db 61 GAGAGGCGCTTTAGCGCTTACGATAAAGACACTCCCGACCGTTGGCAAAACGTTGGCTAAA 120			/codon_start=1 /evidence="not experimental" /product="F18014.2" /protein_id="AAF79456.1" /db_xref="GI:8778448"
Qy 41 AlaValGlyGlyLysVerValGluGluValLysArgHisTyRAspIleLeuValGluAsp 60			/translaton="MSQHLLLLLSLLHLLHKKPISATITTIQKPEAPQFVNSADCPLI DSESDDVVVAKPIFCRRRAVHVMTLDAAYIRGSVAALVSLVLOHSCPENIVPHFVA SASADASSLRATISSSPVYLDFTVYVFNVSRSRLSSISRSALDCPLNARSYLADL LPPCVRVYVLDSDLLIDIAKLAATDLGRDSVLAAPCYCNANFTSYFTSTFWSNPT LSLTADRACYPNTGVMVLDLSEWREGAVTSRIEEMWAMQKMRIRIYELGSLRPFLV FAGLIRPVRNHRNQHGIGDNGFRLGRDLHPGPVSLHWSGKPKPWARLDAGRPCPLD ALWAPYDLILQTPFALDS"
Db 121 GCAGTTGGAAGTAAATCTCGAGAGGAAGTTAAACGTCCTACGACATCCCTCGTTGAAGAT 180			complement(9620..10300) /notes="similar to Zn finger protein gi 6319911"
Qy 61 LeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrlsThrPheGlu----- 78			/codon_start=1
Db 181 CTCATGAACATCGAACAGACTTAGTACCTTGCCTAAATACAAACCGTCGATGTGGA 240			
Qy 79 SerAsnSerArgSerIleAsnAspPheAspThrArg 90			
Db 241 AGTAAATCTAGAGCATCGATGATTTGTTGAGG 276			
AC025808/c			
LOCUS	AC025808		
DEFINITION	Genomic sequence for Arabidopsis thaliana BAC F18014 from chromosome 1, complete sequence.		
ACCESSION	AC025808		
VERSION	AC025808.8		
KEYWORDS	HTG.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	1 (bases 1 to 120977)		
AUTHORS	Shinn,P., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bei,Q., Chin,C., Chiou,J., Choi,E., Conn,L., Conway,A., Gonzales,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S., Mukharasky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vayseberg,M., Yu,G., Federspiel,N.A., Theologis,A. and Ecker,J.R.		
TITLE	Genomic sequence for Arabidopsis thaliana BAC F18014 from chromosome 1		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 120977)		
AUTHORS	Ecker,J.R.		
TITLE	Direct Submission		
JOURNAL	Submitted (15-MAR-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA		
REFERENCE	3 (bases 1 to 120977)		
AUTHORS	Ecker,J.R.		
TITLE	Direct Submission		


```
/evidence=not experimental
/product="F18014.3"
/protein_id="AAF79455.1"
/db_xref="GI:8778447"
/translation="MSDVPSCSGNDTNNDSNFEICLDAQDPVITLGHLPFW
PCLYKWLHLSQKDCPCVKAIEEDRLVPLXGRKSSADPRSKIPLGVNPRPSGQ
RPTAQPPNPHGFAHHHGFGMGGFAAPMASARFGNVTLSAARFGLIPSLNLFHF
GFPDAAHYGAAASGPPHGFSPNPHGGSHMSYQRHTGRGQOQDHILILLIVFV
VFSUFLS"
CDS
join(12523..12577,12686..13374)
/note="hypothetical protein"
/codon_start=1
/evidence=not experimental
/product="F18014.4"
/protein_id="AAF79420.1"
/db_xref="GI:8778412"
/translation="WAIFSTSHLLFISFIATCTTIVSGTTFTLNHCSTIWPGLT
ANGAGGGGFFALASGVSTFTVSPGSGRFWARTYCNFDASGSGCGTDCGSKLKC
AGAGAPATLAEFTIGSGGKNVADFDVSLVDGVYVQMKITPQGGSGDCKTAGCV
SDVNAICPEKLOVTPGSGVAACKSACEAPNKECYCTGATSPATCPPTNYSKIFQA
CPSAYSAYDDASSFTCTNAYEISFCS"
complement(join(14020..14136,14236..14304,14402..14464,
14557..14696,14879..15005,16274..16486))
/note="unknown protein; similar to EST gb|H76493.1"
/codon_start=1
/evidence=not experimental
/product="F18014.5"
/protein_id="AAF79454.1"
/db_xref="GI:8778446"
/translation="MLEAVDSGGVNGGFPQIQSYFGYSCSBEELSILPRHTKVVVTG
NNRTKSVLQGVVKKVAGLGGHMLVLTNGIEVKLQNALSVLEPTGNEEDDDLD
PENTORNGSDTSDTLKPKHKLGRQSSRSRSHKTMSSLSQSSQSGFTPEPNN
KVDLSKLEMPALLNWRHFNLDVAINPNSKQLIIVQHFMSQMDLELVIVGFQA
AKRMKACFKQSKESRNTDLNCIS"
join(17994..19489,18568..18698,18814..18847,19006..19099,
19170..19347,19441..19479,19616..19715,19855..19994)
/note="hypothetical protein"
/codon_start=1
/evidence=not experimental
/product="F18014.6"
/protein_id="AAF79421.1"
/db_xref="GI:8778413"
/translation="MAKTDKLAQFLDSGIYESDEFNWFFLDTVITNRSYTRFKVSPS
AYISRFNNSKOLNHSSESPPKRRKOKNSFHLPSVGEQASNLRHQEARLFLSKAH
ESFLKEILLSITKGLSDDDSSLLNKKCDDDEVSVFELGGVWQAFYITILSPNL
HCDNEGSCNORVQFQPNLVNVEIGEEVEAFNSNRYIIMPRNSCFYMSDLHHIRNL
VPKSEBGNLIVIDPPWENASAHQKSEFVLYMYPTLPQYFLSLKQLAHAEGL
VALWYTRKELISFVEKELFPAGWIKYVATWYLVKVPDGLICDLDLAGEKRSDF
KLLDNQIIMSPIDGFSRKPPIGDIILLKHTPGSQPARCLELPAREMAAGWTSWNEPL
HPQDSRYFLKV"
join(22037..22300,22618..23361)
/note="unknown protein; similar to ESTs gb|AI992723.1,
gb|T41777.1, dbj|AV440980.1, gb|AI997341.1, and
dbj|AV440785.1"
/codon_start=1
/evidence=not experimental
/product="F18014.7"
/protein_id="AAF79422.1"
/db_xref="GI:8778414"
/translation="MTSDGATSTAAAAAAMATRKPSWRENNRRRRRRRAVAA
KIYTLRAQGNVNLPKKDNNEVTKALCSAGWVEEDGTYRKGHKPLPGDMAGSSS
RATPYSHNQSPPLSTFDSPIYSQVSPSSSPSPSPGVDPHNIISTIFPLRNGIIP
SSLPPLRISNAPVTPPVSSPTSRLPKPLTWESFTKQSMMAKQSWTSLNPFYAV
SAPASPTHRRQFHAPATIPCEDSDSTVDSGHWISFOKFAQQQPFPSASWPTPTFN
LVKPAQOOLSNTAAIQIGOSSEFKFNSQVQKPEWGERIHDVAMEDLELTGNGKAH
S"
complement(join(23868..24983,25237..25407))
/note="unknown protein; similar to EST gb|T13858.1,
emb|Z38050.1, and emb|Z38087.1"
/codon_start=1
/evidence=not experimental
/product="F18014.8"
/protein_id="AAF79453.1"
```

```
/db_xref="GI:8778445"
/translation="WAGRRDRSQQLRSGRIATAILLIGIFIGCVCAVLFPYGFNPNSSSS
LKASHLKSSNQVSSGSSSERVAMLKSDFTLSEKNAELKKQVRELTEKURLAEQ
GSDNARKQVLALGTQIKAGPFGTVKSLRTNPTILPDESINPLAKILEEIAVDKEVIV
ALANANVAMLEQVIASIRGITNYLVVALDDYIENLCKENDVAYYKRDPOKVDVT
KTKTGNHVASGLKFRVLREFQLQGYGVLLSDVDIVFLQNPFSHLVYRDSVDSMSDGH
NHTAYGPNDFDEPAMGWARVAHTWRIWVFNSEFFVLRPTIPISELLDRVADRLSKAK
WDOAVFNEELFYPSHPEVLTALHASKRVMDMYEFNNSKVLFTVRKNHKLKKVKPVI
VHVYHPDKLNRMQAVVEFYVNGKQDALDSFPDGSE"
complement(26100..27917)
/note="unknown protein; similar to ESTs gb|T42717.1,
emb|F13886.1, and emb|F13885.1"
/codon_start=1
/evidence=not experimental
/product="F18014.9"
/protein_id="AAF79452.1"
/db_xref="GI:8778444"
/translation="MADPLNGKSFICFSLFSFTLLFTISPLYATESPVIEDVSTDA
VSVSTNREAVLLHNLNLTAVLNDLAKSATPLKKNBISVDDDIQEEKERG
RAKASFVTKYSPFMSERFQTSVAKLNSDATICINLPRDPFGSKYPAIGSKGRVY
VFLRNGDLVIEFFTVDSPTVAMSYSSVFNKNSFVVTGHQNGAVLLHRIHGSNGED
VFNNSVSMHVHGKFDVDSADPTLVLEHVHGRVRYILATLSLGSGLTVLTENRTVYGS
VIPSRLPLVKQRLFLFTESGASDLRSKMKIRETECEGLNHSIARTYVFDAAERSK"
Alignment Scores:
Pred. No.: 1,48e-25 Length: 120977
Score: 343.00 Matches: 69
Percent Similarity: 80.4% Conservative: 5
Best local Similarity: 75.0% Mismatches: 16
Query Match: 52.3% Indels: 2
DB: 15 Gaps: 1
US-10-697-787-2 (1-126) x AC025808 (1-120977)
Qy 1 MetAlaSerAanSerArgSerSerIleSerProTrrPThrPheSerGlnAsnLysMetPhe 20
Db 90486 ATGGCCTCTAGTCTTCTAGCTCGAGCTCTTCTTGGACGCTCTAAGCAAAACGATGTT 90427
Qy 21 GluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrrPheHisValAlaLys 40
Db 90426 GAGAGGGCTTTAGCCGTTTACGATAAAGACACTCCCGACCGTTGGCAAAACGTGCTAA 90367
Qy 41 AlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuValGluAsp 60
Db 90366 CGAGTTGGAAGTAATCTGCGAGAGGAAGTTAAACGCTACTACGACATCTCTGTTGAAG 90307
Qy 61 LeuIleAanIleGluThrGlyArgValProLeuProAanTyrLysThrPheGlu----- 78
Db 90306 CTCATGAACATCGAACAAGACTTAGTACCTTTGCTTAATAACAAACCGTCGATGTGGA 90247
Qy 79 SerAanSerArgSerIleAanAspPheAspThrArg 90
Db 90246 AGTAAATCTAGAGGCATCGATGATTTGATTGAGG 90211
RESULT 8
AC007119/c AC007119 54433 bp DNA linear PLN 27-FEB-2002
LOCUS Arabidopsis thaliana chromosome 2 clone F2G1 map mi238, complete
DEFINITION sequence.
AC007119
VERSION AC007119.6 GI:20198094
KEYWORDS HTG.
SOURCE Arabidopsis thaliana (chale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 54433)
AUTHORS Lin,X., Kaul,S., Shea,T.P., Fujii,C.Y., Shen,M., VanAken,S.E.,
Barnstead,M.E., Mason,T.M., Bowman,C.L., Ronning,C.M.,
Benito,M.-I., Carrera,A.J., Creasy,T.H., Creasy,T.H., Buell,C.R., Town,C.D.,
Nierman,W.C., Fraser,C.M. and Venter,J.C.
JOURNAL Unpublished
```


Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eucosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

1 (bases 1 to 306)
Qu, L. and Gu, H.

TITLE

The MYB Transcription Factor Family in Arabidopsis: A Genome-Wide Cloning and Expression Pattern Analysis

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 306)
Qu, L. and Gu, H.

TITLE

Direct Submission

JOURNAL

Submitted (07-JAN-2004) Life Sciences, National Laboratory of Protein Engineering and Plant Genetic Engineering, Peking University, Beijing 100871, China

FEATURES

Location/Qualifiers

source

1..306
/organism="Arabidopsis thaliana"

gene

1..306
/mol_type="mRNA"

CDS

1..306
/locus_tag="At2g21650"

ORIGIN

1..306
/locus_tag="At2g21650"

Alignment Scores:

Pred. No.: 1..94e-22 Length: 306

Score:

286.00 Matches: 56

Percent Similarity:

74.7% Conservative: 6

Best Local Similarity:

67.5% Mismatches: 19

Query Match:

43.6% Indels: 2

DB:

15 Gaps: 1

US-10-697-787-2 (1-126) x AY519526 (1-306)

QY

1 MetAlaSerAsnSerArgSerSerLe-----SerProTyrThrPheSerGlnAsnLys 18

Db

1 ATGGCATCAGGCTCAATGTCCTTTATGCTCTGGCTCATGCACTGTTAAGCAGAACAA 60

QY

19 MetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTyrHisVal 38

Db

61 GCCTTTGACGCGTGCTAGCAGTCTATGACCAAGACACTCCGACCGTTGGCACAATGTT 120

QY

39 AlaLysAlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuVal 58

Db

121 GCTAGAGCTGTTGGTGGTAAACACAGAGAGCTAGACAGATGACCTTCTAGTT 180

QY

59 GluAspLeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGlu 78

Db

181 CGTGACATCGAAGCATCGAAGTGGTACGTCGACGTCCTGCTACTACAGACTACTACA 240

QY

79 SerAsnSer 81

Db

241 GGAAACAGC 249

REFERENCE

1 (bases 1 to 337)
Yamada,K., Dale,J.M., Huan,V.W., Onodera,C.S., Quach,H.L., Chen,H., Toriumi,M., Wong,C., Yu,G., Yuan,S., Carninci,P., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

JOURNAL

Arabidopsis Open Reading Frame (ORF) Clones Unpublished

REFERENCE

2 (bases 1 to 337)
Yamada,K., Dale,J.M., Huan,V.W., Onodera,C.S., Quach,H.L., Chen,H., Toriumi,M., Wong,C., Yu,G., Yuan,S., Carninci,P., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

TITLE

Direct Submission

JOURNAL

Submitted (16-MAY-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

COMMENT

The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

FEATURES

Location/Qualifiers

source

1..337
/organism="Arabidopsis thaliana"

gene

1..337
/mol_type="mRNA"

CDS

1..337
/chromosome="2"

misc_difference

1..337
/clone="U50297"

3'UTR

1..337
/ecotype="Columbia"

ORIGIN

1..337
/note="This clone is in pUNI 51."

Alignment Scores:

Pred. No.: 2..18e-22 Length: 337

Score:

286.00 Matches: 56

Percent Similarity:

74.7% Conservative: 6

Best Local Similarity:

67.5% Mismatches: 19

Query Match:

43.6% Indels: 2

US-10-697-787-2 (1-126) x AY519526 (1-306)

QY

1 MetAlaSerAsnSerArgSerSerLe-----SerProTyrThrPheSerGlnAsnLys 18

Db

1 ATGGCATCAGGCTCAATGTCCTTTATGCTCTGGCTCATGCACTGTTAAGCAGAACAA 60

QY

19 MetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTyrHisVal 38

Db

61 GCCTTTGACGCGTGCTAGCAGTCTATGACCAAGACACTCCGACCGTTGGCACAATGTT 120

QY

39 AlaLysAlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuVal 58

Db

121 GCTAGAGCTGTTGGTGGTAAACACAGAGAGCTAGACAGATGACCTTCTAGTT 180

QY

59 GluAspLeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGlu 78

Db

181 CGTGACATCGAAGCATCGAAGTGGTACGTCGACGTCCTGCTACTACAGACTACTACA 240

QY

79 SerAsnSer 81

Db

241 GGAAACAGC 249

rosids; eucosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 337)
Yamada,K., Dale,J.M., Huan,V.W., Onodera,C.S., Quach,H.L., Chen,H., Toriumi,M., Wong,C., Yu,G., Yuan,S., Carninci,P., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

Arabidopsis Open Reading Frame (ORF) Clones Unpublished

2 (bases 1 to 337)
Yamada,K., Dale,J.M., Huan,V.W., Onodera,C.S., Quach,H.L., Chen,H., Toriumi,M., Wong,C., Yu,G., Yuan,S., Carninci,P., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

Direct Submission

Submitted (16-MAY-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEN (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Yamada,K., Dale,J.M., Huan,V.W., Onodera,C.S., Quach,H.L., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGEN) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEN) contributed equally to this work as PIs.

Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.

Location/Qualifiers

1..337
/organism="Arabidopsis thaliana"

1..337
/mol_type="mRNA"

1..337
/chromosome="2"

1..337
/ecotype="Columbia"

1..337
/note="This clone is in pUNI 51."

1..337
/gene="At2g21650"

1..337
/gene="At2g21650"

1..337
/codon_start=1

1..337
/evidence=experimental

1..337
/product="putative myb family transcription factor"

1..337
/protein_id="AAP40381.1"

1..337
/db_xref="GI:30793857"

1..337
/translation="MASGSMSSYSGSWTVKQKAFERALAVYDQTPDRWNVARAVGGKTPEAKRQYDLLVRDIESIENGHPFDPYKTTGNSNRGLRDEKRMKMLQ"

1..337
/gene="At2g21650"

1..337
/note="compared to genomic sequence"

1..337
/replace="g"

1..337
/gene="At2g21650"

1..337
/note="compared to genomic sequence"

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eucosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 306)
Qu, L. and Gu, H.

The MYB Transcription Factor Family in Arabidopsis: A Genome-Wide Cloning and Expression Pattern Analysis

Unpublished

2 (bases 1 to 306)
Qu, L. and Gu, H.

Direct Submission

Submitted (07-JAN-2004) Life Sciences, National Laboratory of Protein Engineering and Plant Genetic Engineering, Peking University, Beijing 100871, China

Location/Qualifiers

1..306
/organism="Arabidopsis thaliana"

1..306
/mol_type="mRNA"

1..306
/locus_tag="At2g21650"

1..306
/locus_tag="At2g21650"

1..306
/codon_start=1

1..306
/product="MYB transcription factor"

1..306
/protein_id="AAS09996.1"

1..306
/db_xref="GI:41618982"

1..306
/translation="MASGSMSSYSGSWTVKQKAFERALAVYDQTPDRWNVARAVGGKTPEAKRQYDLLVRDIESIENGHPFDPYKTTGNSNRGLRDEKRMKMLQ"

1..306
/organism="Arabidopsis thaliana"

1..306
/mol_type="mRNA"

1..306
/locus_tag="At2g21650"

1..306
/locus_tag="At2g21650"

1..306
/codon_start=1

1..306
/product="MYB transcription factor"

1..306
/protein_id="AAS09996.1"

1..306
/db_xref="GI:41618982"

1..306
/translation="MASGSMSSYSGSWTVKQKAFERALAVYDQTPDRWNVARAVGGKTPEAKRQYDLLVRDIESIENGHPFDPYKTTGNSNRGLRDEKRMKMLQ"

1..306
/organism="Arabidopsis thaliana"

1..306
/mol_type="mRNA"

1..306
/locus_tag="At2g21650"

DB:	15	Gaps:	1	Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.
US-10-697-787-2 (1-126) x BT008554 (1-337)				
Qy	1	MetAlaSerAsnSerArgSerIle-----SerProTrpThrPheSerGlnAsnLys	18	
Db	1	ATGGCATCAGGCTCAATGCTCTTATGGCTCTGGCTCATGGACTGTTAAGCAGACAAA	60	
Qy	19	MetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnVal	38	
Db	61	GCCTTTGAGCGTCTCTAGCAGCTATGACCAAGACACTCCGGACCGTTGGCAATGTT	120	
Qy	39	AlaLysAlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuVal	58	
Db	121	GCTAGAGCTCTGCTGTTAAACACCAAGAGAGCTAAGAGACAGTATGACCTTCTAGTT	180	
Qy	59	GluAspLeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGlu	78	
Db	181	CGTGACATCGAAAGCATCGAATGGTCAGTACCTCCCTGACTACAAGACTACTACA	240	
Qy	79	SerAsnSer	81	
Db	241	GGAACACAGC	249	
RESULT 14				
BT008698		570 bp	mRNA	linear
LOCUS		Arabidopsis thaliana clone RAFL17-18-C23 (R50297)		putative myb family transcription factor (At2g21650) mRNA, complete cds.
DEFINITION		BT008698		
ACCESSION		BT008698.1	GI:30794123	
VERSION				
KEYWORDS				
SOURCE		Arabidopsis thaliana (thale cress)		
ORGANISM		Arabidopsis thaliana		
REFERENCE		Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsis.		
AUTHORS		1 (bases 1 to 570)		
Yamada,K., Dale,J.M., Hsuan,V.W., Onodera,C.S., Quach,H., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.				
Arabisdopsis Full Length cDNA Clones				
Unpublished				
2 (bases 1 to 570)				
Yamada,K., Dale,J.M., Hsuan,V.W., Onodera,C.S., Quach,H.L., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.				
Submitted (16-MAY-2003)				
Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA				
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.				
The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada,K., Dale,J.M., Hsuan,V.W., Onodera,C.S., Quach,H.L., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A.				
Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC) contributed equally to this work as PIs.				
FEATURES				
source				
1..570				
/organism="Arabidopsis thaliana"				
/mol_type="mRNA"				
/db_xref="taxon:3702"				
/chromosome="2"				
/clone="RAFL17-18-C23 (R50297)"				
/ecotype="Columbia"				
/note="This clone is in a modified pBluescript vector2 (lambda PS) as a BamHI/XhoI insert."				
1..570				
/gene="At2g21650"				
1..8				
/gene="At2g21650"				
9..314				
/gene="At2g21650"				
/codon_start=1				
/evidence=experimental				
/product="putative myb family transcription factor"				
/protein_id="AAP40504.1"				
/db_xref="GI:30794124"				
/translation="MAGSSMSYSGSWTVKONKAFERALAVVDDTPRWHNVARAVGSKTPEAKRQYDLLVRDIESIENGHVPPDYKTTTGNRNRLRDEEXMRSMKLQ"				
misc_difference 221				
/gene="At2g21650"				
/note="compared to genomic sequence"				
/replaces="g"				
315..570				
/gene="At2g21650"				
misc_difference 554				
/gene="At2g21650"				
/note="not present in genomic sequence"				
ORIGIN				
Alignment Scores:				
Pred. No.:	4.12e-22	Length:	570	
Score:	286.00	Matches:	56	
Percent Similarity:	74.7%	Conservative:	6	
Best Local Similarity:	67.5%	Mismatches:	19	
Query Match:	43.6%	Indels:	2	
DB:	15	Gaps:	1	
US-10-697-787-2 (1-126) x BT008698 (1-570)				
Qy	1	MetAlaSerAsnSerArgSerIle-----SerProTrpThrPheSerGlnAsnLys	18	
Db	9	ATGGCATCAGGCTCAATGCTCTTATGGCTCTGGCTCATGGACTGTTAAGCAGACAAA	68	
Qy	19	MetPheGluArgAlaLeuAlaValTyrAspLysAspThrProAspArgTrpHisAsnVal	38	
Db	69	GCCTTTGAGCGTCTCTAGCAGCTATGACCAAGACACTCCGGACCGTTGGCAATGTT	128	
Qy	39	AlaLysAlaValGlyGlyLysThrValGluGluValLysArgHisTyrAspIleLeuVal	58	
Db	129	GCTAGAGCTCTGCTGTTAAACACCAAGAGAGCTAAGAGACAGTATGACCTTCTAGTT	188	
Qy	59	GluAspLeuIleAsnIleGluThrGlyArgValProLeuProAsnTyrLysThrPheGlu	78	
Db	189	CGTGACATCGAAGCATCGAATGGTCAGTACCTCCCTGACTACAAGACTACTACA	248	
Qy	79	SerAsnSer	81	
Db	249	GGAACACAGC	257	
RESULT 15				
CS137904				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
CS137904.1				
GI:72058614				
linear				
DNA				
578 bp				
Sequence 875 from Patent WO2005047516.				
CS137904				
CS137904.1				
GI:72058614				
PAT 09-AUG-2005				

THIS PAGE BLANK (USPTO)